

Rewriting the rules for a post-Cold War world

Next year's World Conference on Science is a unique chance to reassess the dynamics of international scientific cooperation and address the challenges it currently faces. This opportunity must not be squandered.

One of the less anticipated consequences of the ending of the Cold War has been the decrease in political pressure for international collaboration in science. Behind the rhetoric of science as an international language, many of the institutions set up in the aftermath of the Second World War to encourage such collaboration — including Unesco, the United Nations Educational, Scientific and Cultural Organization — were seen by governments partly as a way to reduce tension between power blocs through greater communication and interaction. Even more familiar is the fact that much technical assistance offered to developing nations was intended, explicitly or otherwise, to encourage them to embrace a particular political ideology (whether capitalist or communist).

Thus, although the evaporation of East–West rivalry has resulted in a substantial ‘peace dividend’, there has been a price. The replacement of power-bloc politics by economic competitiveness faces science with an uncomfortable new world. It is therefore all the more timely that Unesco and the International Council for Science (ICSU) should have decided to celebrate the approach of the new millennium with a World Conference on Science, to be held in Budapest next year. This will be attended by government officials and senior scientists responsible for formulating national sciences policies, and is intended to address the challenges that lie ahead for science in both industrialized and developing nations (see page 299). As preparations for the conference enter their final, critical phase, some guidelines can be suggested as to how its overall effectiveness is likely to be enhanced.

Need for concrete objectives

Formulaic statements about the importance of science to the modern world should be strictly limited. Few of those attending the conference will need to be convinced of this message; conversely, those unconvinced will not be present in Budapest, and are unlikely to have their minds changed by newspaper reports, televised speeches or UN declarations. Time spent extolling the virtues of modern science will have a high opportunity cost.

Furthermore, any bold declarations of principle must be accompanied by a commitment to concrete objectives and a realistic sense of the likely availability of the funds needed to put them into practice. Many of those present will remember how the high hopes generated at the last such UN meeting, held in Vienna in 1979, rested on vague commitments to a new, multi-million dollar ‘science and development’ fund — and how little of this money actually materialized.

Equally, while more money for science is still urgently needed in many countries, scientists must avoid falling into the trap of believing that the problems they face can be solved through extra funding alone. Just as important, as international investment banks have been emphasizing in Latin America and elsewhere, is the need to ensure efficient use of funds that are made available. This means less support for science as a vehicle for either personal or national prestige, and increased attention, including the necessary monitoring, to ensure that scientific resources are properly used — for example through the use of peer

review — and effectively and appropriately applied to social needs and problems.

To achieve the latter, the conference should explore the implications of the shift from ‘producer-led’ to ‘user-led’ science policies in both industrialized and developing nations, as expressed, for example, in the British white paper of 1993. This does not mean the disappearance of arguments in favour of support for basic science, for example in helping a country to develop a qualified workforce and a scientifically literate public, or even for its cultural value. But it does increase responsibilities on the research communities who benefit from this support to ensure that, overall, their activities contribute to all these goals.

Safeguarding science as public knowledge

The Budapest conference will also be an important opportunity to re-examine how the principle of equity of access to science is put into practice. Too often, while lip-service is given to the need to provide greater opportunities for women and minority groups in both producing and using science, insufficient efforts are made to challenge the obstacles that stand in the way. And more effective ways need to be found for enhancing the access of researchers in the developing nations to the skills and facilities of those in the industrialized world, for example through short research training courses.

The conference will be well placed to explore and highlight the importance of new forms of collaboration in science. One example, familiar in industrialized countries but less so in developing nations, is the value of regional collaboration in addressing common research priorities and avoiding the costly duplication of research facilities. Other novel possibilities are raised by electronic forms of communication, such as virtual centres or ‘collaboratories’ whose members work together on the web, rather than at physical centres.

Last but not least, it is essential that the conference attempts to formulate principles safeguarding the status of scientific knowledge as a public good in an era of rapidly spreading privatization. The increasing barriers to open scientific communication posed by intellectual property rights and material transfer agreements will be well known to most readers of this journal. So, too, are the problems caused for international collaboration in science by related tensions between developed and developing nations on issues such as access to genetic resources. But the broader implications of each rarely gets a public airing of the type that is on offer in Budapest.

Provided that issues such as those above are addressed in a hard-headed yet imaginative way, both in the remaining months of preparation and during formal and informal negotiations at the Budapest meeting itself, important groundwork will have been laid towards ensuring that future support for science, whether from national governments or international loan agencies, is appropriately and productively used. If not, then any ‘new social contract’ to emerge from the conference, however hard it is waved in the face of governments, will not be worth the paper on which it is written. □



Personal computers spur drive to keep control over copyright

[SAN DIEGO] Grassroots efforts to control the copyright of research articles are springing up across the United States as researchers seek to increase their direct involvement in electronic publishing.

Some universities are studying how to maintain the copyright for their published research, while others have begun the electronic publication of all graduate theses and dissertations. There are also calls for a national repository to hold a portion of the overall copyright of published research.

The initiatives are being driven partly by the rising costs of journals and growing concerns over the consolidation of companies that publish them. But academic librarians say that researchers' personal computers are the real catalysts for change over the control of copyright of scholarly work.

"The machines on the desktop empower faculty and researchers to control scholarly information in a way they have never had before," says Scott Bennett, Yale University's librarian.

One of the main issues of concern is how to maintain the peer-review certification process that makes today's research strong, while allowing electronic publishing to disseminate scholarly work more widely.

This dilemma has given rise to the concept of 'decoupling' the scientific certification process from the publishing of research articles. How this will be accomplished in the

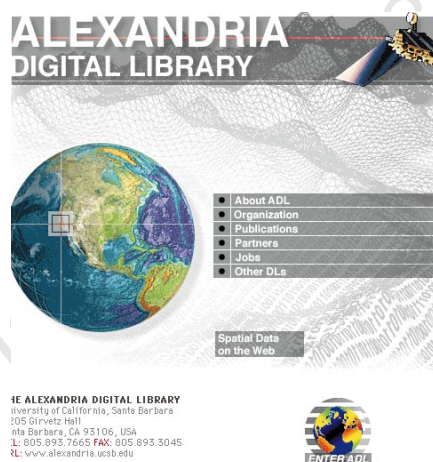
expanding arena of electronic publishing has yet to be determined.

Currently, most journals — and the societies that publish them — require authors to give full copyright to the publication. In other cases, the copyright reverts to the author or authors after it is first published.

There is a growing movement among researchers and universities to fight to retain copyright after the initial publication. "We pay all the costs of generating knowledge, and then we have to pay to buy it back — with an outrageous inflation rate in the cost of journals," says David E. Shulenburg, provost at the University of Kansas.

Last month, at a meeting of the Association of Research Libraries, Shulenburg proposed the setting up of a national repository to hold a portion of the copyright material in academic articles from 90 days after initial journal publication. The National Electronic Article Repository, as he calls it, would facilitate the dissemination of academic work, hopefully keeping costs down while increasing access to knowledge.

The California Institute of Technology (Caltech) in Pasadena has set up a website to collect the opinions of faculty members and students about its plans early next year to assert new controls over the copyright of research. For ten years, Caltech has owned the copyright of all research produced there, but it has not enforced that ownership, says



New world: the growth of digital libraries is leading to a global rethink of copyright issues.

Steven E. Koonin, the university's provost.

After reviewing the discussions, Caltech plans to consider a system whereby the university would retain copyright control after initial publication. This would pit the institution against professional societies, which demand copyright and prohibit subsequent posting of articles on other websites.

To "maximize the distribution of intellectual property at universities, we have to take control of the property right of the work we produce, retain it and use it," says Charles E. Phelps, provost at the University of Rochester in New York.

Nearly two years ago, in the first such action in the country, Virginia Polytechnic Institute and State University took the bold step of requiring that all graduate theses and dissertations be filed electronically at the university library.

There are methods of initially blocking access to research that may be patented, but once such rights have been secured the theses and dissertations are available through the university's library.

Martin Blume, editor-in-chief of the American Physical Society's 11 publications, recently proposed that journals be given licences to publish articles, with the copyright being kept by the authors.

Stefan Von Holtzbrinck, a director of Macmillan Magazines, the publisher of *Nature*, says it is important for publishers of major journals to retain copyright to ensure they can continue to add value to submitted manuscripts through the peer-review process. But he is aware that the scientific community is keen "to ensure that its efforts won't be exploited by a publisher who does not add anything to articles".

Rex Dalton

Germany restores funds to grant agencies

[MUNICH] Germany's coalition government is to make up a shortfall in the budgets of the Max Planck Society (MPS) and the Deutsche Forschungsgemeinschaft (DFG), the university research grants agency, and add an extra five per cent to each in 1999.

Edelgard Bulmahn, the research minister, last week agreed the increases with representatives of Germany's 16 *Länder* (states), which contribute half of the budgets. Assuming parliamentary approval, the MPS will receive a DM90 million (US\$53 million) increase in 1999, bringing its budget to DM1.66 billion. The DFG will receive DM100 million more, raising its budget to DM2.2 billion.

Since 1990, the MPS and the DFG have received an annual five per cent budget increase to build up the science base of eastern Germany. But this year's increase was capped by parliament at 3.5 per cent.

At the time, Hubert Markl, the MPS president, warned that the resulting

shortfall could result in the closure of institutes (see *Nature* 392, 3; 1998). Markl says he is "delighted" by the increase, but remains cautious as the budgets may be capped again by parliament next year.

This is unlikely, however, as the political environment is now more friendly to science. In his inaugural speech to parliament this month, Chancellor Gerhard Schröder reaffirmed a promise to double federal spending on research and education in the next five years.

The MPS wants the money to continue its expansion programme in eastern Germany, called 'Aufbau Ost'. Last week, its senate approved the foundation of an Institute of Ethnological Research in Halle, the twentieth institute in the new *Länder*.

Germany's five new *Länder* now have a similar density of Max Planck institutes to the 11 west German *Länder*. But scientists still work in poor facilities as new buildings are not yet complete.

Quirin Schiermeier

Framework deal boosts life sciences...

[MUNICH] Agreement was finally reached last week on financing for the European Union's long-debated fifth Framework programme of research (FP5). As a result, calls for proposals are expected to be issued in early spring, although money is not likely to reach researchers until the autumn, several months later than planned.

Programmes related to life sciences, training and mobility of researchers and industrial technologies have all emerged with more money than in the fourth Framework programme (FP4). But nuclear research and the EU's Joint Research Centre have fared comparatively badly.

The new programme will last from 1999 until 2002. The delay in its launch was due to the failure of the Council of Ministers and the European Parliament to agree on the level and mechanism of its financing.

Last week the two parties finally agreed to a total of ECU14.9 billion (US\$17.4 billion), an increase of 4.6 per cent over FP4. This includes ECU1.26 billion for nuclear research. The total is almost ECU1 billion more than the council said in February that it was prepared to pay, but is ECU1.74 billion below the budget the parliament wanted.

The agreement was welcomed by UK science minister Lord Sainsbury, who said: "The new budget provides a three per cent increase in real terms over the previous programme. This will be a significant stimulus to the wide ranging collaborative research that Europe needs to promote competitiveness and the quality of life."



Sainsbury: 'stimulus to competitiveness'.

The European Commission, which drew up the programme and is responsible for its execution, had suggested a budget of ECU16.3 billion.

The mechanism of FP5 financing was complicated by Spain's campaign to force member states to continue the programme of structural funds — EU subsidies to poor regions from which Spain greatly benefits — after 2000. The union is debating rule changes for the distribution of structural funds in anticipation of the extension of membership to several central European countries. Spain is using its power of veto on all EU decisions that require unanimity, including FP5, until the issue is resolved in its favour.

As a result, the Council of Ministers could only firmly agree on FP5 financing up to 2000. Parliament was unhappy with the need to reconfirm within 12 months financing of the subsequent three years of the pro-

...but nuclear work and fusion research lose out

[MUNICH] Clear losers in FP5 funding are the nuclear programme and the Joint Research Centre (JRC). The budget for nuclear research, which is organized under a separate Euratom treaty that is yet to be finally approved, has shrunk by six per cent. The JRC, whose eight institutes act as a European centre for science and technology services, sees its budget shrink by more than seven per cent.

The JRC has recently

embarked on a restructuring in response to continued criticisms from politicians that it had become inflexible and of uncertain scientific merit (see *Nature* **395**, 314; 1998).

The European Commission had wanted a zero-growth budget for the centre to allow it to continue on its new track. The final budget is considerably lower than what JRC director Herbert Algeier considers the minimum necessary to keep all its projects going.

Algeier is already discussing with the commission which of the non-nuclear projects will have to be dropped. He says it is already clear that the JRC's nuclear programme will have to abandon the work it does in support of the EU's nuclear fusion programme.

Algeier says that the work on decommissioning nuclear installations should be supported from another part of the union's general budget.

A.A.

gramme. But it conceded when clauses in the 'common position' with the council were reworded to state that parliament would decide jointly about subsequent financing with the council.

As the dust settles, the commission must now fill in the details of FP5's four thematic programmes and its three 'horizontal' programmes, which must be individually approved by the council, with advice from parliament. This less controversial aspect should be complete by February.

Comparisons between FP5 and FP4 in terms of financing and content are difficult because of the radically different way that FP5's programmes have been organized in a bid to concentrate funds more precisely on socially defined problems, known as key actions (see *Nature* **386**, 5; 1997).

But it is clear that life sciences have been given a boost, in line with EU policies to promote biotechnology and to make science serve social needs, such as the improvement of health and the environment.

The thematic programme 'Promoting competitive and sustainable growth', with its focus on industrial technologies, has also won more favour. This is partly a result of pressure from the richer countries which want the research to support the promotion of a European aeronautics industry.

The biggest winner is the horizontal programme 'Improving human potential', which is comparable to FP4's popular 'Training and mobility of researchers' programme in content, and has one third more money.

'Improving human potential' differs from its predecessor in its approach to individual fellowship grants. The postdoctoral fellowships, which allow young scientists to work in a second EU country for up to three years, will remain substantially unchanged, although the number of postdocs that can be funded will increase.

But the heavily oversubscribed three-year PhD fellowships will be modified to pay only travel costs of students wishing to spend 3–12 months of their studies abroad, in a laboratory approved by the commission.

A special scheme will be set up for industry which did not participate in FP4's 'Training and mobility of researchers' to the extent that the commission had hoped. This time industry will be able to define its own projects for three-year PhD fellowships and two-year postdoc fellowships, rather than only being able to accept individuals arriving with approved projects that could be out of line with that industry's mission. The commission will select industrial 'hosts' for researchers.

The commission's idea of selecting hosts is not new. It was introduced in FP3, where it failed because of the commission's slow administrative procedures. Its reintroduction in FP5 is a bit of an experiment, admit officials, but recent administrative improvements mean that fellowships could still be taken up as early as next autumn.

The 'Access to large scale facilities' programme within FP4's 'Training and mobility of researchers' becomes 'Enhancing access to research infrastructure' in FP5. The name change reflects the fact that the programme is able to support research services such as databases or special ecological habitats in addition to large facilities such as synchrotrons.

Scientists are waiting for the details of the programmes to be defined before giving final judgement on the merits of FP5 for basic research. The increase in budget for training researchers is a "very good sign", says Eigil Praestgaard, a professor of chemistry at Roskilde University in Denmark who was a member of the commission's science advisory body, ESTA. "But there remains a lot of grumbling about whether FP5 will be good for basic research when its research areas are so targeted."

Alison Abbott

Radical changes urged in UK curriculum

[LONDON] The British government has responded cautiously to a report on the school science curriculum that suggests that pupils should spend less time memorizing facts and more time pursuing original ideas and developing an awareness of how science works.

The report, published last week by the Nuffield Foundation, is the outcome of a series of nationwide seminars held to review the effectiveness of the school science curriculum. The seminars were held over two years and involved teachers, schools inspectors, educationalists, members of professional science bodies, and the Qualifications and Curriculum Authority, the government body responsible for the school curriculum.

According to the report, pupils should learn how to discriminate between competing scientific claims, how ideas in science change over time, and how to evaluate the strengths and weaknesses of media reports of science. After the age of 14, it suggests, half of the science curriculum time should be spent studying these issues.

Carolyn Swain of the Qualifications and Curriculum Authority says the authority is unlikely to incorporate any radical elements into its review of the school curriculum. This is partly because some of the report's ideas are already part of the curriculum — such as courses in the nature of science, its applications, and science communication — but also to give teachers time to get used to the existing curriculum, which was changed substantially by the previous government.

The overall reaction to the report has been mixed. Most of the recommendations have been endorsed by science bodies, including the Institute of Physics and the Royal Society of Chemistry, which both say they are relaxed at the prospect of a reduction in the factual content of school science.

Colin Osborne, schools and colleges education officer at the society, says there is little in the report that will concern the chemistry community, providing that the proposed new curriculum is not considered a 'soft option' and continues to challenge pupils.

Osborne agrees with the report that the current curriculum does not meet the future needs of most pupils. Similarly, the government's chief scientific adviser, Sir Robert May, says it is "excellent". At the launch of the document, May said that teaching science as a set of facts "left me fairly uninterested in science when I was at high school".

He added: "Many dilemmas in public policy arise because of the belief that science can answer all questions. But the truth is that in many cases we just don't know. Science is not a set of clear answers to well-defined questions."

But employers are less enthusiastic, and

fear that, despite assurances, it may lead to a further decline in standards of assessment that are believed to have been caused by the introduction of the national curriculum a decade ago.

Peter Clark, a senior policy adviser at the Confederation of British Industry, says that, although the proposed reforms have some merit, the above-average performance of British schoolchildren in international education surveys such as the Third International Mathematics and Science Study, suggests that pupils have a better grasp of science than the report implies.

Swayn says that the report's authors seem to have an outdated view of the present curriculum. "The idea that science is taught [wholly] as a discrete set of facts is possibly a hangover from previous versions of the curriculum," she says, adding that any reduction in the facts presently taught must be debated and discussed considerably before — if at all — it is implemented. For example, she says, lessons on the implications of discoveries such as BSE or cloning require a basic understanding of biology.

But the authors deny that their proposals amount to a lowering of standards, another concern of the Confederation of British Industry. The report says that pupils who want to study a subject in more detail — possibly as a step towards a career in science — should be allowed to do so, although Jonathan Osborne, senior lecturer in science

education at King's College London, and one of the report's co-authors, says that these pupils will always be a minority.

He says that the curriculum should aim to meet the needs of the majority of pupils who will not become professional researchers. "This does not amount to 'dumbing down'," says the other co-author, Robin Millar, professor of science education at the University of York.

The curriculum was originally designed as a foundation course for future scientists. But, despite many attempts to make it more relevant to those not intending to pursue careers in science, the authors say that "tension" remains between the aims of promoting scientific literacy and providing the first stage of a training in science.

Some problem areas remain in the curriculum, say the authors, including the fact that pupils find science "uninteresting and alienating", that there is no increase in the number taking science beyond the age of 16, and that many adults have low levels of scientific literacy despite spending about ten years studying science at school.

The authors do not expect their proposals to be implemented in the short term. They also recognize that teachers will need convincing of the need for further change, and that assessment methods must also change, both of which will take time. "It is better to debate these issues now, when there is time to reflect," says Osborne.

EhsanMasood

'Soft option' for women gets sharp riposte

[LONDON] Last week's launch of the Nuffield report (see above) on the future of science education was partly overshadowed by a lively exchange between two panelists on how to get more women to take careers in science.

On the day that Peter Mandelson, cabinet minister for science, launched a poster campaign to attract more teenage girls into science (see right), the authors of the report were asked whether they had addressed this issue. Robin Millar, professor of science education at the University of York, said that, although there was no explicit reference to girls and science, the issue was discussed at the seminars.

He then added that a more "humanistic" curriculum, such as that proposed in the report, would appeal more to girls than the present one. This prompted an angry response from Sir Robert May, the government's chief scientific adviser, who described Millar's comments as "absolute rubbish". It is "pernicious nonsense to say that women like soft things. It's the sort of thing we're trying to combat."

May added: "Countries where women do



Spreading the message: Mandelson launches a bid for more women in UK science.

well in science offer a good middle-class education, and reliable domestic help." The answer, May concluded, is "tax relief for help at home, and better child care".

E.M.

GILES BARNARD

NSF board is urged to act more openly...

[WASHINGTON] The US National Science Board (NSB) is under fire for allegedly failing to adhere to laws intended to make the workings of the US government as open as possible to the public and the press.

John Holmfeld, a freelance journalist and former senior staff member at the House Science Committee, says the board has retreated from its obligations under the 1976 Government in the Sunshine Act. It has defied the act, Holmfeld says, "by conducting almost all of its substantive activities in committees, task forces, and other subgroups that are not open" to the public or the press.

Holmfeld is a veteran observer of the NSB and of the National Science Foundation (NSF), the grant-giving agency it oversees. He recalls that the board initially responded to the act by holding lengthy open sessions in which current and planned NSF projects and programmes were debated.

But in recent years, he says, the real discussion has taken place in closed committees, followed by cursory discussion in open session. For example, he claims that status reports on projects are rarely aired at public sessions of the science board.

He cites a meeting in February 1996, which followed a closed session of the board's Committee for Programs and Plans.

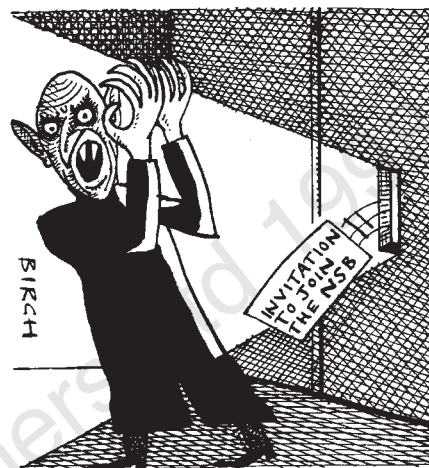
The chairman simply told the full board that "the committee received status reports on the Laser Interferometer Gravitational Wave Observatory, Gemini, Greenbank, Arecibo, the National High Magnetic Field Laboratory, and on NSF management and oversight of the National Center for Atmospheric Research and the academic research fleet".

NSB officials argue that discussions often involve budget and personnel matters that must be confidential. But Holmfeld's view is shared by the small band of reporters who regularly cover NSB meetings.

Some suspect that the committees are used to avoid public discussion of troubled projects, or to avoid public arguments between board members. Earlier this year, for example, the press were kept out for more than an hour while board members haggled over the wording of a public statement resisting the establishment of a National Institute for the Environment within the NSF.

Richard Zare, professor of chemistry at Stanford University and the previous chair of the board, says it needs private discussion for its executive functions. "We have to make decisions that cut programmes, and these couldn't be made in open session," he says.

The board would "get in huge trouble" if it discussed the NSF budget proposal openly



before its publication. "We had more open discussion time when I was chair than before," says Zare. Rita Colwell, the director of the National Science Foundation, agrees. "The board is much more open now than when I was a member in the 1980s," she says.

The previous director of the NSF, Neal Lane, who is now President Bill Clinton's science adviser, concedes that there is room for improvement in the balance between closed and open business at board meetings. "When I was there my sense was that it was uneven: we didn't always have the balance right," he says.

Eamon Kelly, president of Tulane University and the new chairman of the board, said last week that he was not familiar with Holmfeld's complaint, but pledged to look into it.

The impact of the US 'sunshine' laws has receded since the initial burst of illumination accompanying their enactment in 1976. At some agencies, public scrutiny serves to maintain openness. At the National Institutes of Health, for example, drug companies and disease advocacy groups work in close alliance to maximize public discussion at advisory panel meetings.

Similarly, at the Department of Energy, bodies such as the High Energy Physics Advisory Panel have a long tradition of fairly open discussion. (As purely advisory committees with no executive functions, these fall under the Federal Advisory Committee Act, a different statute from that governing the NSB.)

But the NSB has never been challenged on its interpretation of the sunshine laws. Dan Greenberg, the former editor of *Science and Government Report*, took legal action to open up meetings of scientists advising the US president. He says that the NSB is not influential enough to warrant such a challenge.

Holmfeld disagrees. It is "unfortunate and counterproductive" that the board does not conduct more of its business in public, he says, particularly as Lane is "talking about the importance of the public understanding of science".

Colin Macilwain

...as head seeks to spread the word on grants

[WASHINGTON] Rita Colwell, the new director of the National Science Foundation (NSF), said last week that many good scientists "feel cut out" by the agency because they are unfamiliar with its grant procedures.

Colwell told a meeting of the National Science Board, the agency's governing body, that some people who fail to get NSF grants "really don't know the culture" of the agency. "They have never sat on a peer-review committee," she said.

She pledges to inform scientists better about how the agency works. But Colwell denies that her comments imply that the NSF, which funds most US non-biomedical university research, is funding the wrong research. "I take a large and broad view" of whom the agency should support, she says.

Colwell says her concern

arose during discussions with the principal investigator of a group that had what she considered to be an excellent proposal rejected by the agency.

Colwell's comments mark something of a departure for the agency, which prides itself on distributing grants purely on the basis of scientific merit, without regard to the circumstances of those submitting applications. They suggest that she intends to respond to critics of the agency – including congressional representatives of small states – who complain that too much of its money goes to a few elite research universities.

Colwell is the first woman and the second life scientist to lead the agency. She is expected to emphasize opening up the NSF to a broader constituency of scientists and to support

more multi-disciplinary work. "This place can change," she told the science board. "It can react very quickly, and I'm delighted about that."

Joe Bordogna, the deputy director, is leading a task force to review the existing programmes targeting women and minorities, which cost \$110 million. The agency is likely to replace some or all of these with an approach to encourage diversity across all of its work.

Agency officials hope the review will protect affirmative action programmes from legal challenge – US courts have recently deemed programmes based on race to be illegal – while raising their effectiveness. The NSF is believed to be studying an approach to help minority and women postdoctoral researchers, for example, by tracking the progress of such researchers that it sponsors.

C.M.

French supplies law threatens research

[PARIS] French researchers are warning that last summer's efforts by the Ministry of Economy and Finance to tighten up the implementation of a law on the ordering of equipment and materials by public institutions risks bringing public research in the country to a halt.

Already the laboratories of the Centre National de la Recherche Scientifique (CNRS) have been prevented from paying some of their suppliers. Laboratories belonging to other public research agencies will soon find themselves in the same situation.

The law requires every public institution in France to seek bids for national orders worth more than FF300,000 (US\$53,000). Companies submitting the winning bid would become the sole source of supply to that institution for the length of the contract, which will be at least one year.

Until now, public research organizations have been able to use an annual exemption from these requirements to give them time to modify their purchasing policies.

Some progress has already been achieved in this direction, as the CNRS has reduced the number of suppliers for certain products from 65 to three or four. But the Ministry of Economy and Finance feels there should be only one supplier. The issue has become even more sensitive since a recent reform increased the penalties for preferential treat-

ment of contractors, including those covered by the law on public purchasing.

CNRS officials say research institutions have not been able to adapt to the law on public purchasing because they cannot predict their needs on a long-term basis. Hélène Richard-Foy, director of research at CNRS's Laboratory of Eukaryote Molecular Biology in Toulouse, says the law is incompatible with the way research is carried out: "For antibodies, for example, it is impossible to rely on a single supplier as antibodies recognizing the same protein but coming from different sources will not have the same reactivity or effectiveness according to the experimental approach being used."

François Coulier, a researcher at the biomedical research agency INSERM, says: "How can we react immediately to the appearance of a new reactive agent produced by suppliers who have been eliminated by the purchasing strategy?" He points out that "the 400 small high-technology companies that Claude Allègre [the research minister] wants to see created will have no chance of becoming suppliers to public research institutions."

The current arrangements end at the end of December, but CNRS officials say the normal procedure for seeking bids for supply contracts takes between six and eight months.

Last month, several researchers at

INSERM's Institute of Cancer and Immunology in Marseilles sent a letter to Allègre with a petition signed by more than 1,300 scientists. They asked Allègre to approach the finance minister, Dominique Strauss-Kahn, to try to resolve the situation. Allègre replied that he had been trying to do this for several months. "My attempts have not yet achieved results, but a positive outcome remains possible," he said. Nothing has been heard from Allègre since then, so it now seems to be up to the prime minister, Lionel Jospin, to resolve the issue.

CNRS officials have already run up bills with suppliers with which they do not have contracts. But the agency's financial controller, who reports direct to the finance ministry, has been refusing to honour these commitments, resulting in an impasse.

These steps by CNRS seem to have precipitated government action. According to an official in Jospin's office, a new text addressing the situation is being checked by lawyers before being submitted this week to the state council for approval.

The official says the text confirms the research institutes' exemption, and so is unlikely to please the finance minister. It would give laboratories flexibility in purchasing policies, and has been drafted in a way that is compatible with planned reforms in public purchasing policy.

Eric Glover

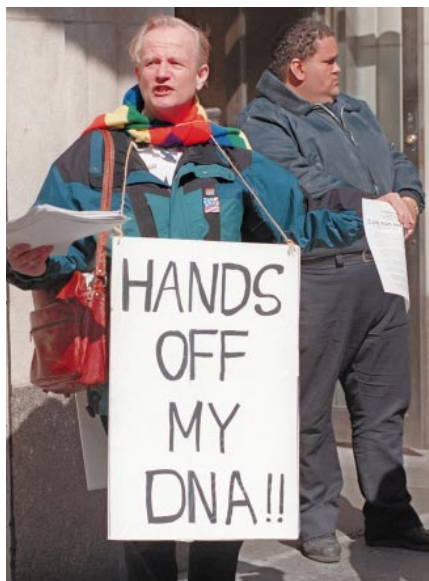
Human genome declaration looks set for United Nations approval

[PARIS] The Universal Declaration on the Human Genome and Human Rights, adopted last year by the United Nations Educational, Scientific and Cultural Organization (Unesco), is set to be adopted by the United Nations itself at its fifty-third general assembly on 10 December.

Supporters of the initiative hope it will give greater political prominence to the human rights issues that, they say, are raised by advances in genetics.

This week the new government of Germany said it would support the text. Germany was the only country to have opposed adoption of the text in last year's vote at Unesco, and as late as last week, at a meeting of the UN commission on human rights, said it had not yet decided on whether to endorse the text.

The approval of the text by the commission, combined with German support, means that it now seems certain to be adopted at the general assembly — which coincides with the fiftieth anniversary of the UN's Universal Declaration of Human Rights. The text is to be presented as a 'procedural resolution', which means that its approval will not require a vote but simply



Lobbying pays off: UN declaration will include condemnation of human cloning for reproduction.

an 'endorsement' of the existing Unesco text.

The declaration, which has been strongly promoted by France, has already attracted 30 sponsors, including Japan, India and

Canada. Dozens of other countries, including the United States and United Kingdom, will support the resolution.

The declaration commits signatories to opposing genetic reductionism and its consequences. It endorses the principle of scientific freedom, although stipulating that this should never be given priority over human rights.

But it avoids specific provisions on medical practices, on the grounds that translating broad principles into legislation is best left to individual states, acting in the light of their own traditions — although Unesco bowed to pressure last year and added a condemnation of human cloning for reproductive purposes and germ-line gene therapy (see *Nature* 388, 508; 1997).

The text will be non-binding on member states. But, like existing UN declarations, it will provide grounds for groups or individuals claiming to have been discriminated against to challenge laws or practices that are contrary its provisions. It also provides a benchmark for countries planning to draft bioethics legislation — such as Japan and many developing countries.

Declan Butler

South African official quits ministry post

[CAPE TOWN] The top government official in South Africa's Department of Arts, Culture, Science and Technology resigned last week following reported disputes with his minister, Lionel Mtshali, and one of his deputies, Musa Xulu, over the use of government funds.

Roger Jardine, the department's respected director-general, is said to have been worried by public perceptions that Xulu and Mtshali were attempting to use the department to advance the political interests of the Inkatha Freedom Party (IFP), which partners the African National Congress (ANC) in the coalition government.

In particular, the department has been criticized for spending almost a million rand (US\$178,000) on a feast for 10,000 guests to be held next month at the unveiling of a new monument to the Battle of Blood River, where the Voortrekkers and Zulus clashed in 1838 (the monument itself is costing R2.5 million).

Most of those invited are local residents and, with an election due next May, this has been interpreted as an attempt to curry favour with the IFP's political base in rural Kwazulu-Natal.

Mtshali was unavailable for comment this week. But his press officer, Frans Basson, confirmed that Mtshali and Jardine held a meeting earlier this month after which Jardine wrote to the minister requesting that his five-year contract be terminated on 31 December — 13 months early.



Mtshali: concern over use of funds.

Jardine, 33, a medical physicist who was the youngest director-general of a government department when he was appointed in February 1995, issued a statement saying that "his reasons for leaving were both personal and professional".

But he is understood to have had an uneasy relationship with Mtshali since the latter took over as minister from Ben Ngubane just over two years ago. Jardine was formerly head of the ANC's science and technology division at its headquarters at Shell House in Johannesburg (see *Nature* 373, 651; 1995).

The ministerial portfolio is one of three in the coalition government allocated to the IFP. But Mtshali, a former schoolmaster turned inspector of schools, appears less pragmatic than his predecessor, who enjoyed a good working relationship with Jardine before moving to Kwazulu-Natal to become provincial premier.

Jardine's is the third departure of a director-general from central government in the past six months. Olive Shisana, who left the health ministry to take up a senior post at the World Health Organization in Geneva, last week openly blamed the Minister of Health, Nkosazana Zuma, for spending R14 million on a controversial AIDS play, 'Sarafina', with-

out her approval or following correct tender procedures (see *Nature* 386, 6; 1997).

The government's anti-corruption unit issued summonses on several parties earlier this month, including both Zuma and Shisana, who are liable for the recovery of outstanding funds for this project.

Jardine's departure has been lamented in research circles. Rolf Stump, deputy vice-chancellor at Stellenbosch University and former president of the Human Sciences Research Council, said that Jardine's spearheading of the transformation of the country's science and technology system "has been one of the most successful systematic transformations in our society".

Deputy president Thabo Mbeki is expected to win next year's elections and to make government more streamlined. The Arts, Culture, Science and Technology ministry has been identified by the presidential review commission as a candidate for dissolution.

If dissolved, its functions are likely either to be incorporated into the Department of Education, or to be split, with science and technology being moved to the Department of Trade and Industry under the highly-regarded minister Alec Erwin.

Ironically, however, the department might survive as a result of the need for an apparently innocuous cabinet portfolio which can be allocated to the IFP, which appears set to remain in a coalition government with the ANC despite questions about its administrative abilities. **Michael Cherry**

Budget cuts mean tough times ahead for Brazilian researchers

[SÃO PAULO] The Brazilian government has slightly reduced the severity of cuts being imposed on science in a general restriction on public expenditure designed to cushion the impact of the international financial crisis on the country's economy.

But the cuts, introduced primarily to meet the terms of a loan from the International Monetary Fund (see *Nature* 395, 831; 1998), remain severe, and have already generated widespread protests from the scientific community. Many scientists are lobbying their members of Congress, who will vote on the detailed budget proposal on 15 December.

The rectors of the nine universities that perform most of the country's research issued a public protest during a recent meeting in Belo Horizonte. Jacques Marcovitch, rector of the University of São Paulo, the country's largest, is worried because federal finance is an important source of funds for graduate research.

"We risk losing the investment made in recent years, and now face the possibility of

compromising basic activities," says Sergio Henrique Ferreira, president of the Brazilian Society for the Development of Science (SBPC).

Overall, the budget of the Ministry of Science and Technology will be cut by 18.7 per cent, or 169.8 million réals (US \$142 million), less than the originally envisaged cut of 907 million réals. Other research-funding ministries will be affected to varying degrees. For example, the budget of the Ministry of Agriculture will shrink by 17 per cent, and the Ministry of Health will lose about 6.6 per cent.

The science ministry has not yet decided which areas will be affected. Ministry officials say only that existing fellowships will be funded, as these are considered to be equivalent to researchers' salaries, which must not be interrupted. But there are likely to be major cuts in trips to attend scientific meetings abroad, and in the financing of such meetings inside Brazil.

Of greater concern to scientists are the prospects for grants to specific research

programmes, such as the competitive Support Programme to Centres of Excellency (Pronex), which is funded by the ministry of science to reward some of the best research groups in the country.

It has already been decided which projects will be funded under Pronex 1 and 2, but researchers are still waiting for some of the promised grants. There is also worry about the next phase, Pronex 3.

Etelvino Bechara, a chemistry researcher at the University of São Paulo, warns that science in the region will be hurt by any cut in support for the National Council for Scientific and Technological Development, the main federal funding agency. "The difficulties of importing research material will probably increase," he adds.

Many researchers, particularly in areas lacking strong regional science funds, are suffering from cuts imposed earlier this year. "We have several projects paralysed in our university," says Lucio Flavio Moreira, a physiologist at the Federal University of Rio Grande do Norte. **Ricardo Bonalume Neto**

'Science summit' sets ambitious agenda

[PARIS] Momentum gathering behind plans for a global 'science summit' next June appears to be fulfilling the organizers' hopes that it could be the most important international meeting on the relationship between science and politics for 20 years.

The World Conference on Science for the 21st Century is to be held in the Hungarian capital, Budapest. It is being organized jointly by the United Nations Educational, Scientific and Cultural Organization (Unesco) and the International Council for Science, until recently known as the International Council for Scientific Unions. (The council is keeping its old acronym, ICSU.)

Continued debate over the effectiveness of Unesco, which the United States still declines to rejoin after withdrawing in the mid-1980s, gave rise to considerable scepticism about plans for the conference when it was proposed last year by the agency's director-general, Federico Mayor.

Concerns remain over a tight planning schedule, a shortage of funds and an ambitious agenda. The latter includes a comprehensive overview of the achievements of modern science, and what one Unesco official describes as a desire to establish "new international guidelines for science policy".

Set against these, however, is a growing awareness that the ending of the Cold War, whose dynamics determined much of the past 50 years' political thinking about international cooperation in science, makes the timing of such a meeting appropriate.

Several of those who plan to attend say privately that, although unconvinced that the meeting will produce a substantive outcome, it could be a valuable opportunity to advance negotiations already under way, and perhaps to float new ideas.

High level delegations of scientists and government officials are expected from more than 120 countries. Those leading delegations or due to give keynote speeches include Bruce Alberts, the president of the US National Academy of Sciences, Israel Vargas, science minister of Brazil and president of the Third World Academy of Sciences, and the British government's chief scientific adviser, Sir Robert May.

It seems likely that the US delegation will be led by Neal Lane, recently appointed as the president's science adviser in Washington, and include Rita Colwell, the new director of the National Science Foundation. If confirmed, their presence would be a particularly significant move in the light of the continued hostility to Unesco in the United States.

"The scepticism has not gone away; but the momentum is certainly building," says one official at the US academy. "The United States, for example, is planning to send a strong delegation of seven scientists and



Flying the flag: the headquarters of Unesco, which is seeking a 'new social contract' for science.

three or four government representatives, and is taking the conference very seriously."

The partnership between Unesco and ICSU, both based in Paris, as organizers of the conference is based on ICSU's role as an informal science advisory body to the UN organization. Three-quarters of Unesco's International Science Advisory Board, where much of the planning for the conference is taking place, are ICSU members.

The scientific community's central role in the planning process stands in sharp contrast to the last such meeting, the United Nations Conference for Science and Technology in Development (UNCSTD) which took place in Vienna in 1979. That was dominated by governments and aid agencies, and Unesco's marginalization still rankles.

"We felt it was important this time that scientists should have a central voice in the planning of the conference," says Istvan Lang, a former foreign secretary at the Hungarian Academy of Sciences who is chairman of the local organizing committee. Hungary is matching the US\$500,000 that Unesco has committed to the meeting with a similar sum to cover the extra costs of holding the meeting in Budapest rather than Paris.

The conference will run from 26 June to 1

July. At the end of the meeting, delegates are expected to endorse two declarations embodying guidelines for follow-up activities, both of which it is hoped will eventually be endorsed by the United Nations.

One, known as the 'Declaration on science', will emphasize the need for a political commitment to science and the resolution of tensions between science and society. The other, 'Science agenda — framework for action' will set out ways to encourage greater scientific cooperation in addressing national and global needs and problems.

The two organizing bodies admit to slightly different goals. ICSU sees the conference as a way of getting the value of science across to politicians. "We would like to try to increase the understanding of science on the part of policy-makers," says Jean-François Stuyck-Taillandier, ICSU executive director.

Unesco also wants to cast science in a positive light, but stresses the need to use science effectively and to challenge the cultural and political barriers to achieving this. "We need to develop a new social contract for science," says Maurizio Iaccarino, Unesco's assistant director-general for science.

Iaccarino is keen to see a rewriting of the informal contract, expressed in its best known form by President Harry Truman's science adviser Vannevar Bush, under which scientists were left largely to set their own research agendas. "Bush's ideas have been used by many natural scientists, but are no longer appropriate," says Iaccarino. "We need now a new commitment of politicians to science, and of scientists to society. Our idea is to put scientists and politicians together to discuss these issues."

Iaccarino says he hopes practical ideas will emerge from the conference. One could be enhanced international commitment to postgraduate training in specific fields, particularly for researchers from developing nations. This might be provided in the form of short research training courses in the industrialized nations.

David Dickson

Indian anger at promotion 'caste system'

[NEW DELHI] A decision announced last week by the Indian government to limit the opportunities for promotion for scientists working outside the 'strategic' departments of space, atomic energy and defence has provoked an angry reaction among affected researchers.

Researchers working in these three areas will continue to enjoy a fast-track promotional system, but those in the eight 'non-strategic' government departments, such as biotechnology, industrial research, and oceanography, will have to wait between

three and five years before being assessed for promotion to the next grade.

The announcement was made by the Department of Personnel after the rejection of a suggestion from a national pay commission that scientists in all government agencies should enjoy a uniform pay structure and promotional opportunities. The Joint Action Forum, which represents researchers in the eight non-strategic scientific departments, says the government is legitimizing a caste system in Indian science.

K. S. Jayaraman

France's INSERM to share control of labs with universities

[PARIS] A framework agreement has been signed between France's Ministry of Education and Research, the Conference of University Presidents and the national medical research agency INSERM on the relationship between the agency and the university system. Under the agreement, responsibility for individual INSERM institutes will be shared with universities.

The agreement has been reached without the knowledge of researchers, who have been confronted with a *fait accompli*. Many are concerned about the absence of a clear definition of the status of INSERM. "This is proof of a desire to dismantle the agency just when its representative bodies are about to be renewed," says Claude Mawas, head of one INSERM research team.

The agreement gives universities joint responsibility for INSERM units, indicating that 'joint research units,' which must be linked to local university laboratories, should be the norm. The heads of such units will be nominated by the university and INSERM.

There will be no requirement to share resources within the unit, and doctoral students will be supported through the

university. A protest petition addressed to prime minister Lionel Jospin is already gathering signatures. "This is the end of the autonomy of INSERM to fix its research activities and promote an agency-wide strategy," says Mawas.

All change at US science funding committees

[WASHINGTON] The fate of the budgets of US science agencies will be in fresh hands after next January. A reshuffle of the chairs of appropriations subcommittees in the House of Representatives will follow the elevation of Bob Livingston (Republican, Louisiana) from chairmanship of the main appropriations committee to speaker of the House. James Walsh (Republican, New York) will replace Jerry Lewis (Republican, California) as chair of the subcommittee for veterans' affairs, housing and urban development and independent agencies, which has responsibility for the space agency NASA and the National Science Foundation.

Chairmanship of the energy and water subcommittee, which control funds for the Department of Energy, passes from Joseph McDade (Republican, Pennsylvania) to Ron Packard (Republican, California). But John Porter, a champion of biomedical research, will stay put at the labour, health and education

subcommittee, which controls funding for the National Institutes of Health. Livingston himself is replaced as full committee chair by Bill Young (Republican, Florida).

Brézin will stay on at French science agency

[PARIS] Edouard Brézin has been reappointed by France's Council of Ministers as president of the Centre National de la Recherche Scientifique (CNRS). The action had already been referred to by Claude Allègre, the research minister, in newspaper interviews earlier in the month.

Allègre has asked Brézin, a physicist, to draft new statutes for the CNRS "with the freedom to choose the form of consultation which he judges to be useful". In delegating this task to Brézin, rather than the director general of the agency, Catherine Bréchignac, the minister has broken with tradition.

Germany ups cash for university buildings

[MUNICH] Germany's new research minister, Edelgard Bulmahn, has announced her intention to add DM200 million (US\$118 million) to the federal share of the university building programme, which has been controversially capped at DM1.8 billion per

year since 1995. In her first speech to parliament as a minister, Bulmahn “invited” the *Länder* (state) governments, which share the cost of the building programme equally with the federal government, to discuss increasing investment in university infrastructure.

But her proposal falls far short of the DM4.7 billion which the Wissenschaftsrat (German science council) claims is the minimum required next year to save universities from “serious damage to their research infrastructure”.

Ethics group advises against embryo ban

[PARIS] Calls for a ban on European Commission funding of research involving the destruction of human embryos were opposed this week by the European Group on Ethics in Science and New Technologies, the body that advises the commission on ethical matters.

The European Parliament is to vote on an amendment that would outlaw such funding (see *Nature* 396, 105; 1998). But in an ‘opinion’ released on Monday (23 November), the group argues that embryo research is ethically and legally a matter for member states, given that attitudes depend on various cultural and religious traditions.

The group also argues that the commission lacks the means to adequately inform the public of the ethical and scientific issues raised by advances in the life sciences.

Space station's first module takes off

[WASHINGTON] Construction of the International Space Station, the largest engineering project ever attempted in orbit, began last week with the launch of the Zarya module from Kazakhstan. Built with US funding, the Russian module will provide power and propulsion for the station during its first few months in orbit.

Next week a US space shuttle will carry up the station's second piece, a docking module with multiple ports. Russia's service module, the early station's nerve centre and living quarters, is to follow next summer. As this critical third piece is behind schedule because of lack of funding from Russia, the US space agency NASA is grudgingly paying \$60 million to see it finished.

UK institutes merge in efficiency drive

[LONDON] Britain's Natural Environment Research Council has approved the merger of its freshwater and terrestrial ecology

research institutes in Dorset on to the nearby site of a former atomic energy research centre. All 60 of both institutes' staff are expected to move in after the centre has been refurbished in 18 months.

The merger will strengthen interdisciplinary research, according to Mike Roberts, director of the Institute of Terrestrial Ecology. The merger was suggested under the previous government's ‘prior options’ review of the effectiveness of government research laboratories.

Plea to revive search for civilizations in space

[WASHINGTON] Fans of the Search for Extraterrestrial Intelligence (SETI) programme, including Stephen Jay Gould, the Harvard evolutionary biologist, science fiction writer Arthur C. Clarke and 30 celebrated writers, journalists and scientists, have written to President Bill Clinton asking him to restore its funding.

NASA support for SETI, which seeks to detect radio signals produced by civilizations in space, began in 1992 but ended abruptly the following year, after being subjected to mockery in Congress. The petition says that SETI is a scientific bargain at \$12 million a year, and that with further public support it could “capture the public imagination”.



Back on track: the rebirth of human genetics in China

David Dickson

Concern about the growing interest of foreign companies and research institutes in China's wealth of genetic resources has helped the country's geneticists to cement political support for their own research efforts. In the process of doing so, the country appears to have successfully cast off the anti-Mendelian beliefs of its early communist era.

The eyes of many geneticists were focused on Beijing last summer, after a decision to hold the 18th International Congress on Genetics there had generated world-wide controversy over China's declared intention to use genetic screening as part of its birth control strategy.

Ironically, however, the strenuous efforts by Chinese geneticists to arrange for the meeting to take place in Beijing — and to oppose efforts to shift it to another location in the face of boycott threats — had been partly prompted by a very different concern, the need to secure full backing for human genetics in political circles.

Three factors have helped China's geneticists to press their case. One is the fact that Western biotechnology and pharmaceutical companies are increasingly turning to genetics as a potential source of new diagnostic techniques and treatments, prompting researchers in China to argue that its own industry needs to move in the same direction, and requires a domestic base in fundamental genetics to be able to do so.

Second is the fact that many of the Chinese postgraduates and researchers working abroad, particularly in the United States, and whom China is keen to attract back to help create a modern, science-based economy,

work in genetics. This perhaps reflects the solid grounding in basic mathematical skills provided by China's education system.

Third, the keen interest shown by foreign companies in gaining access to genetic data from the Chinese population — in particular from the many close-knit and genetically distinct ethnic minorities — has been skilfully used to emphasize that in this population, China possesses a potentially valuable economic resource. But, it is argued, this resource can only be exploited in the country's interest if a domestic research community is sufficiently developed to carry out this task.

Set against these pressures for change, however, has been the continuing legacy of three historical events: the dominance of communist genetics in the 1950s by the anti-Mendelism of the Russian agronomist Trofim Lysenko, the Great Leap Forward at the end of that decade, and the Cultural Revolution of the late 1960s and early 1970s.

The weight of history

Each of these events has taken a heavy toll on genetics research. Yet before the communist takeover in 1949, genetics had been relatively developed in China, characterized for example by the pioneering work of researchers such as the *Drosophila* geneticist Ruqi Li, who had worked with T. H. Morgan at Columbia University, New York.

All this changed with the arrival of Russian technical 'advisers' in the early 1950s, bringing with them Lysenko's ideas that Mendelian genetics should be rejected as capitalist ideology, and banished from school textbooks, university courses and all research programmes.

"Geneticists had a terrible time," says Huanming Yang, professor of genetics at the Chinese Academy of Sciences, and director of the academy's recently opened Human Genome Centre in Beijing. "For about two generations of students, the subject was not taught at all. 'Gene' was a bad word, a slogan of the bourgeoisie."

The expulsion of Russian advisers as a result of the split between Russia and China in the late 1950s brought a brief respite, during which some of those who still remembered their earlier genetics training were able to resume their research; indeed the academy's Institute of Genetics was founded in 1959.

But the relief did not last long. First came the Great Leap Forward, a broad movement focusing on projects of immediate practical application. Researchers were pressured into directing their energies into achieving an immediate — and, as it turned out, unattainable — growth in industrial and agricultural production.

There are stories from this period of futile attempts imposed on plant researchers, and still guided by Lysenkoist ideas, to persuade

On other pages | [Ethnic diversity creates opportunities: p304](#) | [Controversy brings moral issues to the fore: p306](#)

crops to grow in totally unsuitable locations. Widespread food shortages resulted when these efforts proved unsuccessful.

Faced with such failures, Mao Zedong turned to a new strategy, the infamous Cultural Revolution. At the insistence of Mao's Red Guards, university professors and researchers were sent to work on the land to be 're-educated', many university departments were closed, and work on human genetics again came to a virtual halt.

Few classically-trained geneticists were able to sustain this triple blow to their discipline. Many left to work in the United States, never to return. Others died in the fields, or moved into other areas of research, such as plant physiology. When the US virologist Howard Temin visited China in 1977, he concluded that genetics played virtually no role in Chinese biology.

One of the few to have survived was the geneticist C. C. Tan (Tan Jiazhen), a student of Li's who was for many years director of the Genetics Institute at Fudan University in Shanghai. Tan, now in his 90s, has been a chief architect of the revival of genetics over the past 20 years — and was the president of last summer's international congress in Beijing.

A survivor's victory

Tan started his research career in the early 1930s, publishing a series of papers, specialising in the genetics of *Drosophila* and ladybirds, between 1932 and 1949. The influence of Lysenko meant that he then published nothing until 1963, with a further ten-year gap after 1966 during the Cultural Revolution. It was only a personal friendship with Mao that allowed him to keep the Shanghai institute open as a lonely outpost of Mendelism.

Today Tan is one of the most fervent supporters of the idea that China needs to establish its position as a leading nation in modern human genetics, and to use its skills in this field to meet the country's social needs and to provide the basis of a competitive presence in the global economy.

Speaking at the opening of the Beijing meeting, for example, Tan described how genetics was an essential tool for addressing two of the most pressing problems facing the country: the need to feed its 1.2 billion population — a staggering 22 per cent of the world's population — and to develop the medical techniques necessary to meet its health needs.

Partly under Tan's leadership, genetics in China has progressed steadily over the past 20 years. Achievements have been particularly marked in plant genetics, where the potential of recombinant DNA techniques has been rapidly absorbed, for example, in work on sequencing the rice genome.

There has also been progress in human genetics. Important 'disease genes' have been

identified through studies of the Chinese population — most recently those implicated in liver cancer. Other achievements include the identification of novel mutations associated with haemophilia and non-insulin dependent diabetes, and China now has 30 laboratories involved in human genome research.

According to many Chinese geneticists, however, it has still been something of an uphill task to convince political authorities to support this work. Lysenko cast a long shadow, and the idea of genes as significant factors in helping to determine the fate of individuals has been difficult to accept by many of those brought up to believe such determinancy to be a myth.

Ironically, a Western threat of a different kind has recently become one of the most powerful weapons geneticists have been able to use to press their case for more support — namely that multinational companies have shown intense interest in exploiting China's wealth of genetic resources.

This wealth is based on China's astonish-

ing ethnic diversity — and the fact that many of its ethnic groups have retained their biological identity over centuries. The distinct characteristics of these groups, aided by administrative records that often go back over many centuries, have become a gold-mine for population geneticists. Researchers are keen to use family relationships to identify the genetic basis of — and if possible to identify the gene or genes responsible for — particularly prominent physiological symptoms (see box).

Western interest

"This huge population provides a unique resource for isolating disease genes through linkage analysis in large numbers of families," says Zhu Chen, one of the country's leading geneticists and the director of the Shanghai Human Genome Research Centre — the second of the country's two main centres for genome research.

Inevitably, Western researchers are showing great interest in the potential of this vast and unique genetic resource. Numerous

Mining a rich seam of genetic diversity

"The populations of Finland and Iceland have become valuable resources for geneticists because of their relative isolation. In China, we have several Finlands and several Icelands," says Lin He, a neuropsychiatric geneticist at the Shanghai Research Centre of Life Sciences.

With the rapid growth of interest in population genetics in recent years, China has become a gold-mine for such studies. Although 93 per cent of Chinese are from a single ethnic group, the Han, the country also has 55 minority groups. Many of them — such as the Tibetans or the Muslim Uyghurs in Xinjiang — live in remote border regions.

Through accidents of geography and history, each group has maintained its cultural — and genetic — identity over hundreds, if not thousands, of years. The result is a unique resource that offers vast potential for the study of the genetic basis of human evolution and diseases.

Chinese researchers, backed by foreign colleagues, are developing the skills to exploit these opportunities. The evolutionary studies are already playing a key role in the international Human Genome Diversity Project. And various groups — watched closely by Western pharmaceutical companies — are studying the genetic basis of diseases from asthma to cancer in minority groups in which the incidence is high.

Many Western researchers are also making full use of the genetic variety available to them to study. But as China's economic development leads to the geographical dispersal of minority groups

and more intermarriages, the window of opportunity is closing fast.

"With more and more intermingling between different genetic isolates, the conservation of genome samples is an urgent task," says Zhu Chen, head of the Shanghai Human Genome Centre.

In response, one of the first goals adopted by China's Human Genome Project has been to preserve at least 50 sample genomes from each minority group — as well as genome samples from groups of Han in both the north and south of the country.

Much of this work is concentrated in the south-western province of Yunnan, whose proximity to the foothills of the Himalayas has left it with a wealth of plant, animal and human genetic variety.

Kunming, the capital of the province which boasts 25 ethnic groups, has become a national focal point for the collection of samples through the Institute of Medical Biology of the Chinese Academy of Medical Sciences. "Our key aim is to preserve samples from China's minority groups," says Chu Jiayou, director of the institute.

Parallel work on conserving genome samples from animals and humans is taking place at the Kunming Institute of Biology, part of the Chinese Academy of Sciences. "This is one of the biodiversity hotspots of the world," says Ya-Ping Zhang, head of its laboratory of cellular and molecular evolution, which has collected more than 10,000 genetic samples in the past two years alone. "We have to make a big effort to preserve what would be very difficult to regain if it was lost."

joint studies between Western university-based research groups and Chinese counterparts have been set up to investigate the genetic factors predisposing to diseases ranging from cancer to schizophrenia.

The overall level of genetic diversity has also helped to persuade China to become an active participant in the Human Genome Diversity Project being spearheaded by L. Luca Cavalli-Sforza, of the Stanford Medical School. Already, China's population geneticists are making their impact felt domestically through, for example, recent studies that challenge the widely-held belief in a separate, Asian, origin for modern man (see *Proceedings of the National Academy of Sciences* 95, 11501–11503; 1998).

Equally inevitably, however, this interest has been shared by biotechnology and pharmaceutical companies in the West, eager to add genetic discoveries to their battery of diagnostic and therapeutic techniques for treating common Western diseases.

In some cases, collaboration has been built up in the full glare of publicity. Several major Western pharmaceutical companies, for example, including familiar names such as Smith Kline Beecham and Glaxo Wellcome, have forged alliances with Chinese research teams.

Smouldering nationalism

Other efforts have been conducted more quietly. It was only on reading a press release on a different topic distributed in October 1995 by the California-based biotechnology company Sequana, for example (since absorbed into Axys), that the Chinese authorities became aware that the company had been using samples collected in China to identify a candidate asthma gene.

Such information has fanned the embers of a smouldering Chinese nationalism that has for several centuries (and, to judge from historical events such as the Opium Wars, perhaps with good reason) been wary of exploitation by foreigners. It has also been put to good use by geneticists in seeking greater political support in their search for research funding.

In particular, much of a recent surge in support for research in genome sequencing in China can be traced back to a letter written by Tan to the country's leaders last year. Tan complained that China already had "a serious gene drain problem", and commented on the country's possible fate if it failed to invest in genome research — and also to patent its genetic discoveries.

"A gene is a kind of wealth," wrote Tan. "If China does not get its own gene patents, then in the next century its biotechnology industry — and in particular its pharmaceutical industry — will be like 'the Admiral of the Northern Fleet who saw all his ships capsize and sink beneath the waves'" — a reference to a familiar event in Chinese history.



Human value: ethnic diversity has become an important national resource.

Tan's appeal did not fall on deaf ears, and struck a chord in a government newly committed to expanding its science-based industries — including those in the life sciences. According to an article last summer in a weekly newspaper in Beijing, China's President, Jiang Zemin, whose brother died of liver cancer, one of the diseases being closely studied, responded that "if we are not concerned about a danger when it is far away, we will certainly have greater worries when it is near".

The most immediate effect of these cries of alarm has been to persuade the government to introduce strict regulations requiring formal approval to be given to all research projects involving genetic material done through collaboration between domestic and foreign research groups. Adequate provisions must be made for the sharing of any royalties that accrue from eventual patents (see *Nature* 395, 5; 1998).

Privately, representatives of Western companies and academic research groups are worried about the red tape and bureaucratic delays, which are likely to escalate. According to reports from the US embassy in Beijing, for example, the law left some US researchers working on joint projects unable to take samples out of China to confirm the

analyses by their Chinese research partners.

But Chinese officials and their scientific advisers argue that the regulations should, in practice, facilitate future collaboration by removing the political suspicions that surround them at present. "The important thing is that we want to work with foreign partners in a way that both sides are protected," says Chen. "We do not want to push people away."

If tighter regulation of access to genetic resources by foreigners — and, by implication, enhanced access by Chinese geneticists — has been one fall-out from the nationalistic complaints, another has been the increased public and political profile that the debate has given to the potential value of genome research in general.

One of the leading figures responsible for raising this profile has been Boqin Qiang, first vice-president of the Chinese Academy of Medical Sciences, and a chief adviser to the government's State Science and Technology Commission. Qiang warned the government that the increased patenting by foreign companies of genes discovered in China could lead to a loss of control over medical and health technologies in the next century. He used this argument to persuade the commis-

sion to make research into the separation, cloning, structure and function of genes a top priority in the country's so-called '863 plan' intended to promote high technology.

One immediate outcome was the decision earlier this year to launch major initiatives in human genomics, including the creation of two human genome centres — one in Beijing, headed by Qiang, and the other in Shanghai, headed by Chen. The centres have a proposed budget of 250 million yuan (US\$30 million) over the next three years (see *Nature* 394, 109; 1998).

The enthusiasm for the new genetics that characterizes the government's programme is reflected in the private sector. Some see a business opportunity in supplying the gene-based diagnostic techniques that the Chinese will need as genetic screening becomes integrated into its medical system.

Commercial opportunities

Foreign companies, many based in the United States and Europe, have already identified the growing market in China. But China itself is eager to promote the creation of small Western-style biotech 'start-up' companies, equally keen to exploit the new market opportunities. Many hope to do so by attracting back Chinese geneticists who have received their research training in the West.

Other companies are already seeking to

provide the equipment and services that will be needed for domestic genome sequencing programmes. Shanghai Genecore Biotechnologies, for example, has been set up as a joint venture by the Applied Systems Division of Perkin Elmer, Axys Pharmaceutical and SiniWest Holdings of San Diego, California.

The main goal of the company is to provide large-scale automated DNA sequencing and sequence analysis services for research institutions throughout China. "We want to provide sequencing services for research laboratories, and have, for example, already been co-operating with the Shanghai Cancer Institute in sequencing mutations that may be linked to liver cancer," says Jumin Wang, senior medical adviser to the company.

Words of caution can be heard among all this enthusiasm. Some warn, for example, that China may be trying to run before it can walk, and should be concentrating on building up basic education and research skills in genetics, as well as competence in core information handling technologies, before entering into the financially risky waters of global competition.

"Chinese research, administrative and business leaders should heed the lessons from Asia's current financial crisis, namely that a 'copycat' economy without a substantial research and development effort and

without its own intellectual property is unsustainable," says Matthew Huang, who is both a marketing manager with Pangea Systems Inc., in Oakland, California, and deputy director of the Shanghai Human Genome Research Centre (see page 307).

Others see a different type of danger, namely that a growing enthusiasm among undergraduates and young researchers for genetics-based explanations of biological processes risks overshadowing other approaches that remain equally relevant to the prevention and treatment of disease. "Now perhaps there is an excessive tendency to say that everything is genetic, everything comes from the genome," says Chen of the Shanghai Centre. "That is not good, especially among some young researchers."

Finally, there are those who argue that the apparent excesses of the so-called 'eugenics' law reflect an eagerness among some senior geneticists, after being shunned for so long by the political establishment, to re-establish their social status by demonstrating the potential contribution of medical genetics to human health and welfare — sometimes leading them to make exaggerated claims as to what genetics can achieve.

Mao Xin, a psychiatric geneticist at the West China University of Medical Sciences currently working at Britain's Institute of Cancer Research, points out that Chinese clinical geneticists are considerably more prepared to share genetic information with relatives and third parties than their western colleagues. The explanation, he suggests, may be that "they hold strong social views regarding the use of genetic information, which they think is important not only to the patient but to society".

Awareness of the controversy that can be generated by such issues among Western colleagues has resulted in an acknowledgement among a growing number of Chinese geneticists of the need to address the economic, social and legal implications of their subject (see left). It has also linked discussion about the development of genetics in China directly to broader questions about political transparency, economic liberalization and respect for individual human rights — central, for example, to discussions of 'informed consent'.

Each is a vast issue with many controversial aspects that Chinese geneticists are somewhat reluctant to discuss. But despite the uncertainty that they create, Qiang and others point out that a country with a wealth of genetic resources now has a government — including the science minister, Zsu Lilan — committed to developing those industries capable of turning these resources to competitive advantage, and actively to support a scientific community gearing itself up to provide the necessary science base. They remain convinced that human genetics in China has returned to the mainstream. □

A crash course in ethical behaviour

China's geneticists have learnt the hard way that, if they wish to increase their role in the international research community, they must come to terms with the social and moral concerns over the applications of genetics that confront their Western colleagues.

The learning experience was the controversy that led up to last summer's International Congress of Genetics in Beijing over certain clauses in China's Maternal and Infant Health Care Law of 1994. These contained what some non-Chinese geneticists criticized as excessively eugenic goals in proposing systematic abortion, based on genetic screening, to reduce the incidence of mental disease.

Much of the heat was taken out of the debate by a detailed discussion at the congress itself, as well as by clarifications and modifications of the law by the Chinese authorities (see *Nature* 394, 711; 1998). But one of the debate's legacies has been awareness of the way that some foreign geneticists had threatened to suspend cooperation with Chinese researchers if action was not taken.

"The law has been an obstacle to international collaboration," says Huanming Yang, director of the Human Genome Centre in Beijing. "Some foreign colleagues have

said to me — only half jokingly — 'if we work with you, will the results be used to implement this law?'"

Yang himself has been a tireless advocate of the need to raise awareness among Chinese geneticists of the economic, legal and social issues (ELSI) raised by their research. "Education [of clinical geneticists] in ELSI has not been seriously started [in China]," he says. "It is a long-term issue; but we have to start now."

A desire to ensure that such issues are placed at the centre of genetic practices in countries such as China has been a key motivation behind the declaration on the human genome adopted last year by the United Nations Educational, Scientific and Cultural Organization (see *Nature* 390, 110; 1997). Yang is keen to see the principles contained in this declaration embedded in Chinese practice.

Zhu Chen, director of the counterpart Human Genome Centre in Shanghai, agrees with Yang's concerns, stressing that the success of China's efforts in genome research depends heavily on the level of international cooperation that can be achieved. "Indeed, we welcome suggestions and criticism from our Western colleagues," says Chen. "Anything that we feel we should not be doing here should be changed."

Growing pains of a genomics industry

Sir—Your report on Chinese genomics ventures is interesting and informative (*Nature* 394, 601–602; 1998; see also Briefing in this issue, page 303). Although this is clearly an exciting time for such activities in China, one could argue that, at present, the proposed commercialization of genomics may do more harm than good.

A business cannot be built out of a vacuum. Some of the initial requirements are a sophisticated financial infrastructure to support high-risk investment, a pharmaceutical industry strong enough to support start-up companies and form critical alliances, and a social culture that can tolerate the 'trial and error' nature of such start-ups. Also needed are an effective legal system for protecting intellectual property, the ability to develop new technology, and an advanced science base capable of making the discoveries that are essential to success.

China does not meet these conditions. For example, the research and development budget of an average US pharmaceutical company is about 20 per cent of revenue. But in China, most companies do not spend any money on research, as the drugs they produce are 'copies' of drugs developed

elsewhere. Furthermore, the 1993 Chinese patent law implies that genes cannot be patented, while in genomics research there are only a few complementary DNA sequencing programmes, and positional cloning, genomic mapping and sequencing have yet to start.

Hastily launching a genomics industry in this environment is premature. But the process has already begun. Genomics is being over-promoted by academics, with promises that will not be kept, and without the necessary substantial budgetary support from the government. In industry, some ventures have already been set up with little understanding of the marketing or financial factors involved.

It may take many years for a start-up to reach profitability, and the premature loss of backing may curtail an otherwise successful venture. The recent setback to genomics ventures backed by real-estate concerns in Shanghai and Beijing illustrate this. Such quasi-commercial ventures do not only damage the reputation of the scientists and research institutions involved, they also confuse policy makers and private citizens, evoking an old Chinese saying that if you go too fast, you may not reach your destination.

Genomics does, however, present a significant opportunity for China to advance its biotechnology research. And the Chinese research community has shown much maturity in recent conferences and meetings, acknowledging that it will take consistent and persistent efforts to reach its goals. Some of the problems mentioned here reflect a common growing pain for the Chinese Human Genome Project which is still in its infancy.

Chinese leaders in research, administration and business should heed the lessons of the financial crisis in Asia, namely that a copy-cat economy is unsustainable without committed research and development efforts and the protection of intellectual property. The best way to achieve success in the industrialization of genomics is through long-term, focused investment, and sustained efforts to build a strong, internationally competitive research programme. For this goal, China needs dedicated genome scientists more than scientists-turned-entrepreneurs.

G. Matthew Huang

Pangea Systems Inc., 1999 Harrison Street, Suite 1100, Oakland, California 94612, USA, and National Human Genome Center, Shanghai, China

Combat climate change by reducing fertility

Sir—By treating future population as fixed, David Victor¹ and M. I. Hoffert *et al.*² frame the problem of stabilizing atmospheric CO₂ as a wholly technological one. But policy-makers need to be aware of all the factors that can significantly reduce the energy demands they highlight.

The mid-range population projection employed² is chiefly dependent on the questionable assumption of replacement-level fertility rate (2.1 births per woman) in the long run³. If policy options that improve social welfare shift that rate by less than half a birth per woman (to 1.7) — a rate above that in many industrialized countries — the projected population drops by 18% in 2050 and by 46% in 2100 (ref. 4). The total energy demand will drop by the same amount (by 5 TW and 20 TW respectively), considerably easing CO₂ stabilization².

Improvements in reproductive health, education and gender equality are desirable and also tend to reduce fertility³. It should interest policy-makers that accelerating them will also help stabilize climate.

Stuart R. Gaffin

Environmental Defense Fund, New York 10010, USA

Brian C. O'Neill

Brown University, Providence, Rhode Island 02912, USA

1. Victor, D. G. *Nature* 395, 837–838 (1998).
2. Hoffert, M. I. *et al.* *Nature* 395, 881–884 (1998).
3. Bongaarts, J. *Science* 282, 419–420 (1998).
4. United Nations. *World Population Projections to 2150* (UN, 1998).

Radical reform needed to aid junior scientists

Sir—The call to limit the number of PhD students at least acknowledges that there is a growing glut of junior scientists (*Nature* 395, 101 & 103; 1998). It is, however, a futile gesture that fails to recognize that the glut is merely a symptom. The heart of the problem is how we recognize and reward the success of group leaders.

A group leader's apparent productivity is measured by the number and impact of the publications of their group. With more publications they can acquire more grant money. With this funding they can hire additional junior scientists to generate more publications. The oversupply and devaluation of postdoctoral scientists is the inevitable consequence. Symptoms of the same disease that drives the glut are nepotism and honorary authorship; blocking the

grants or publications of competitors; the abuse of power; and even fraud.

How can we break the vicious cycle? One approach may be to change how we measure a group leader's performance. The funding system should measure the 'effectiveness' of group leaders by dividing their research performance by the effective size of their groups. Funding according to effectiveness would shift the selection pressure on groups from size to efficiency, and would allow small, efficient groups and even individuals to remain viable.

The current funding system permits group leaders to hold multiple grants which are used to employ junior investigators for specific projects. The system sanctions the erosion of the intellectual freedom of junior scientists and exposes them to abuses of power. These problems would be avoided if junior scientists were funded directly and could use this funding to apply to join existing research groups. A group's funding would then be determined, in part, by the number of junior scientists it could attract.

Young scientists deserve a better system. I urge the scientific community to consider these underlying problems seriously and to act soon on workable solutions.

Gerard Vassiliou

University of Ottawa Heart Institute, 40 Ruskin Street, Ottawa, Ontario K1Y 4W7, Canada

Cryptic clues revealed

Andrew Cossins

Darwinian evolution is often thought of as incremental and progressive, although the fossil record indicates step-wise change in morphological characteristics. An analysis of mutant fruitflies now provides a mechanism for rapid morphological change, with implications for evolutionary theory.

Of all biological processes, the generation of a multicellular, differentiated organism from an egg is surely the most complex. Cells proliferate rapidly to create sheets and blocks, which are then folded and shaped to form the tissues that differentiate into characteristic embryonic structures¹. Coupling proliferation with cell migration involves a high degree of precision in specifying both the fates of individual cells and their spatial relationships with each other — even minor changes in this complex ballet can result in deformed bodily structures. Yet developmental defects are rare, suggesting that a powerful suite of control mechanisms exists. Now, on page 336 of this issue, Rutherford and Lindquist² report that, in the fruitfly *Drosophila melanogaster*, mutations in an unusual heat-shock protein, Hsp90, can elicit a range of specific and substantial developmental defects. Unusually, affected flies survive to adulthood and are fertile, allowing these defects to be reinforced and sculpted by selective breeding.

Heat-shock proteins have a long history. They were discovered as a series of microscopic 'puffs' in the giant chromosomes of salivary glands from heat-shocked fruitflies³. These puffs represent transcription of a small set of genes, resulting in the preferential synthesis of the corresponding gene products, the heat-shock proteins, which range in relative molecular mass (M_r) from 15,000 to over 100,000 (ref. 4). Many of these, including Hsp70 (M_r , 70,000), confer stability on cellular proteins. They help to restore the native folding of proteins that have been disrupted by high temperature⁵. Also called chaperones, they have been found in all organisms studied so far and they are now known to protect against other stresses including oxygen starvation, free-radical damage and even some chronic degenerative diseases. So they are more correctly described as 'stress proteins'.

Hsp90 is one of the more abundant chaperones, but its function is less well understood. At normal temperatures it binds to a specific set of proteins, most of which regulate cellular proliferation and embryonic development⁶. These signalling proteins form complex webs of molecular switches

that allow signals both within and between cells to be transduced into responses, and the coordination of these responses⁷. Hsp90 is thought to stabilize these inherently unstable proteins so that they can respond to stimulation by upstream components of the signalling system⁸. But after heat shock (and, presumably, other stresses), Hsp90 can increase the rate at which stress-damaged proteins are reactivated. So, heat shock can divert Hsp90 from its usual protection of specific signalling proteins.

Several laboratories have isolated fruitflies with mutations in Hsp90. Mutations in both copies of the gene are lethal, so the mutation can be maintained only in heterozygous flies — those that have copies of both normal and mutant Hsp90 genes. Rutherford and Lindquist² noticed that, unlike normal flies, a small proportion (1–3%) of the mutants were deformed, with a range of malformed structures including legs, wings, antennae, eyes and bristle patterns (Fig. 1). These flies survived to adulthood and were fertile, properties not generally observed with mutations in other genes.

In breeding experiments using eye or

wing malformations as markers, Rutherford and Lindquist found that each deformity was heritable. However, these defects did not show simple Mendelian genetics, suggesting that they arise from many genetic determinants (that is, they are polygenic). The genetic background of the flies was important in defining the exact kind of malformation, because crossing the mutant into different lines of fly produced different, but characteristic, defects. Selective breeding of flies with eye deformities led, over only a few generations, to deformity in 80–90% of the progeny, showing that pre-existing genetic determinants of this effect could be accumulated into more potent combinations by artificial selection. Surprisingly, none of these highly selected flies retained the Hsp90 mutation — enrichment of the genetic factors that produced the malformed flies allowed the eye defect to be expressed in the absence of the original enabling mutation.

There are several reasons for invoking Hsp90 in the production of these malformed flies. First, identical malformations were produced when different Hsp90 mutants were crossed into a single line of normal flies. Also, by crossing two flies, each with different Hsp90 mutations, more severe and more frequent malformations were produced. Second, the proportion of deformed flies was increased at high rearing temperatures. Third, and most convincingly, a similar range of malformations was produced when normal adults were fed with geldanamycin, a specific inhibitor of Hsp90.

Because malformations occur in specific structures, it is unlikely that mutation of Hsp90 simply blocks the signalling functions of its normal target proteins. Instead, Rutherford and Lindquist suggest a more subtle effect in which the delicate balance of the signalling pathways that underpin

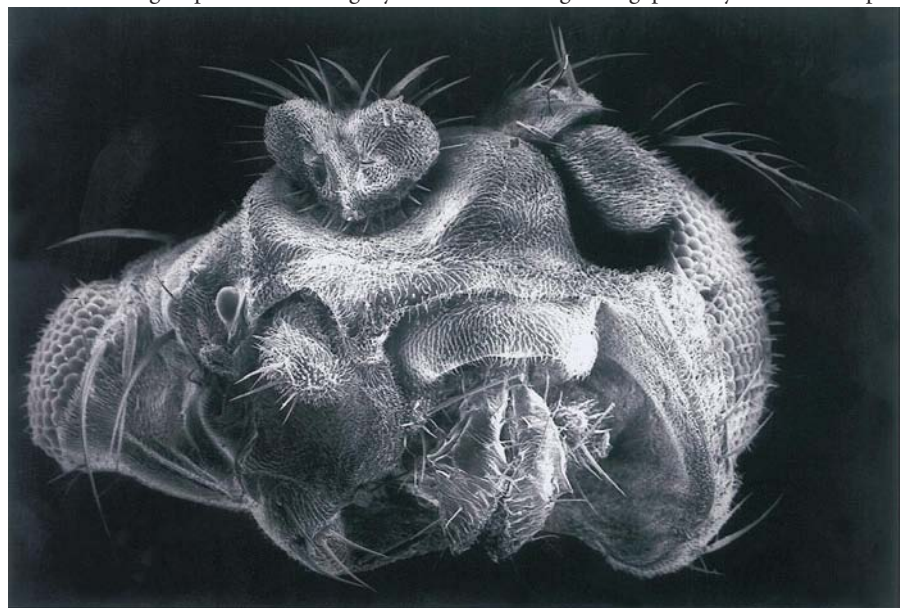


Figure 1 Large and rapid morphological change. Fruitflies with mutations in Hsp90 show a range of developmental defects in, for example, the eye.

CHRISTINA M. ULANE, UNIV. CHICAGO CANCER CENTER ELECTRON MICROSCOPY FACILITY

NATURE

100 YEARS AGO

Some ten years ago a French missionary started the systematic rearing of two kinds of spiders for their web, and the *Board of Trade Journal* states that a spider web factory is now in successful operation at Chalais-Meudon, near Paris, where ropes are made of spider web intended for balloons for the French military aeronautical section. The spiders are arranged in groups of twelve above a reel, upon which the threads are wound. It is by no means easy work for the spiders, for they are not released until they have furnished from 30 to 40 yards of thread each. The web is washed, and thus freed of the outer reddish and sticky cover. Eight of the washed threads are then taken together, and of this, rather stronger yarn cords are woven, which are stronger and much lighter than cords of silk of the same thickness. These spider web ropes are very much more expensive than silk ones, but it is hoped to reduce their cost somewhat in the future.

From *Nature* 24 November 1898.

50 YEARS AGO

A statement on the forest policy for Uganda was recently issued by His Excellency the Governor... The main points are as follows: 1. (i) To reserve ... sufficient land (either already forested or capable of afforestation) to maintain climatic conditions suitable for agriculture, to preserve water supplies, to provide forest produce for agricultural, industrial and domestic purposes, and to maintain soil stability in areas where land is liable to deterioration if put to other uses. (ii) To manage this forest estate to obtain the best returns on its capital value and the expenses of management ... (iii) To foster, by education and propaganda, a real understanding among the people of Uganda of the value of forests to them and their descendants. (iv) To encourage and assist the practice of sound forestry by local authorities and private enterprise; and to educate selected Africans in technical forestry. ... The supply in perpetuity of the many forms of forest produce required to satisfy the wants of the people of Uganda ... must be assured by the reservation, preservation, development and management of the minimum area of forest land needed for this purpose.

From *Nature* 27 November 1948.

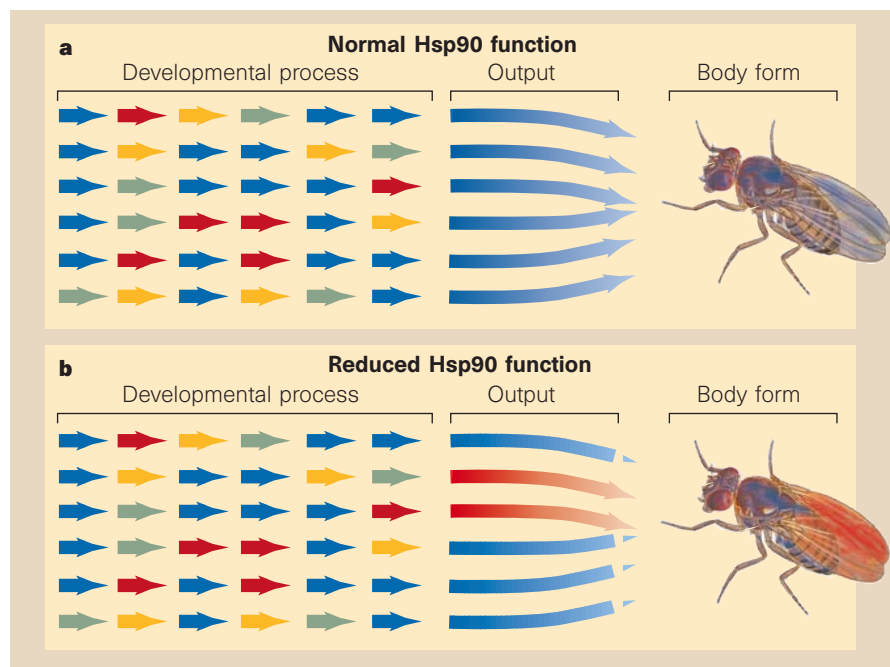


Figure 2 Effect of Hsp90 on development. The arrows represent proteins from the many regulatory pathways involved in an organism's development. The genetic background contains variants that can influence Hsp90-dependent traits (green, yellow and red arrows), as well as those that are independent of Hsp90 (blue arrows). **a**, When Hsp90 function is normal, the signalling pathways have a normal output (long blue arrows) resulting in a normal body form. Under these conditions the genetic variants are cryptic and are selectively neutral. **b**, When Hsp90 function is impaired, some signalling pathways become sensitized; that is, the function of Hsp90-dependent proteins is altered. The pathway may now be affected by genetic variation in other component proteins, the number of affected components depending on the severity of Hsp90 impairment; green, yellow and red components producing effects with progressively greater Hsp90 impairment. These variants may disrupt the overall output of the signalling pathway (long red arrows). Subtle changes in these interlinked signalling pathways can disrupt the normal integration of development leading to phenotypic abnormalities.

development and morphogenesis is altered. This alteration renders the developmental process sensitive to genetically based variations in the target proteins bound by Hsp90, as well as in other signalling components. Provided enough of these weaknesses are co-expressed, the resulting fly has an altered morphology, the nature of which depends on the variants present (Fig. 2).

Although this mechanism is entirely novel, it does not transgress conventional genetics or evolutionary theory. Instead, it offers an unusual twist on the way that genetically based variation affects the form of the resulting body. By sustaining the readiness of particular proteins to participate in signalling reactions, Hsp90 acts as a capacitor against which genetic variation can build up without any effect on the expressed characteristics. Reduced Hsp90 surveillance — resulting from a loss-of-function mutation, or by diversion into the stress-induced chaperone role — allows this hidden variation to be discharged. Uniquely, Hsp90 links the response to environmental stress with control of development, providing a means by which increased environmental variability might be countered by more variety in the form and function of an

organism. Although most of these modifications would be deleterious, others may be favoured by selection. This mechanism provides, for the first time, a means by which changes in body form can be substantial and step-wise rather than incremental and progressive. The fossil record contains many examples of apparently rapid changes in body form interspersed with periods of relative stability⁹. Whether or not this new genetic mechanism accounts for this controversial idea will undoubtedly be the subject of much argument. □

Andrew Cossins is in the School of Biological Sciences, University of Liverpool, Derby Building, Brownlow Street, Liverpool L69 3BX, UK.
e-mail: cossins@liverpool.ac.uk

1. Wolpert, L. et al. *Principles of Development* (Oxford Univ. Press, 1998).
2. Rutherford, S. L. & Lindquist, S. *Nature* **396**, 336–342 (1998).
3. Ritossa, F. *Experientia* **18**, 571–573 (1962).
4. Tissieres, A., Mitchell, H. K. & Tracey, U. M. *J. Mol. Biol.* **24**, 389–398 (1974).
5. Morimoto, R., Tissieres, A. & Georgopoulos, C. *The Biology of Heat Shock Proteins and Molecular Chaperones* (Cold Spring Harbor Laboratory Press, Plainview, New York, 1994).
6. Nathan, D. F., Vos, M. H. & Lindquist, S. *Proc. Natl Acad. Sci. USA* **94**, 12949–12956 (1997).
7. Pelech, S. L. *Curr. Biol.* **3**, 513–515 (1993).
8. van der Straten, A., Rommel, C., Dickson, B. & Hafen, E. *EMBO J.* **16**, 1961–1969 (1997).
9. Gould, S. J. & Eldredge, N. *Nature* **366**, 223–227 (1993).

Eruption dynamics

Rumbling geysers (and volcanoes)

Bruce R. Julian

Scientists are as susceptible as anyone to the spectacular and beautiful eruptions of geysers. Volcanologists, though, have a special interest in geysers as laboratories where they can study processes that might also occur in volcanoes¹.

Writing in the *Journal of Geophysical Research*², Kedar *et al.* report the most detailed observations yet of vibrations produced by Old Faithful geyser in Yellowstone National Park, made with the goal of understanding the phenomenon of volcanic tremor. These new observations confirm that the main source of Old Faithful's noises is the collapse of bubbles in its water column, the same process that makes a teakettle sing. On the other hand, they show no evidence of acoustic organ-pipe resonance in the geyser's water column, a process central to most models of volcanic tremor.

Volcanic tremor is a characteristic ground vibration that usually precedes eruptions and almost always accompanies them. It is often associated with evidence of subsurface magma flow, but it is not understood how tremor is generated. Proposed processes include crack propagation, flow-induced vibration (as in the action of vocal cords), unstable supersonic flow, bubble collapse and turbulence^{3,4}. In the absence of physical understanding, using tremor in volcano monitoring remains largely empirical. If such understanding could be achieved, we might be able to deduce from seismic observations such things as flow rates and directions, and fluid properties such as gas



Figure 1 Hot and steamy science. Old Faithful is a natural laboratory for volcano seismologists.

content and viscosity, which are of obvious relevance to understanding, and perhaps predicting, volcanic activity.

A great difficulty in studying tremor, of course, is the impossibility of making observations within the magmatic systems of active volcanoes. Many studies have therefore centred on vibrations around geysers, in the hope that they are analogous to tremor.

The relatively predictable Old Faithful (Fig. 1) has been the most studied subject. Under the watchful eyes of park rangers anxious to preserve its health, this most famous of all geysers has over the years been eavesdropped on by seismometers and probed by sampling vessels, thermometers, pressure sensors and even a television camera⁵. Now Kedar and his colleagues² have deployed digital seismometers around the geyser, and simultaneously recorded temperatures, pressures and sounds in the top few metres of the water column.

The ground vibration around Old Faithful varies systematically during an eruption cycle, beginning with a quiet period after an eruption, then increasing for about ten minutes to a high level that persists until the next eruption. The build-up is not monotonic, but shows variations that are similar from cycle to cycle, and can be mimicked by a simple plumbing model with an appropriately chosen conduit geometry.

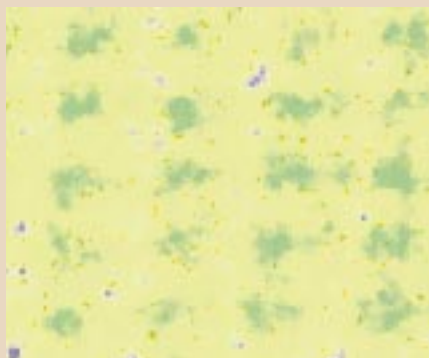
On the new digital seismograms, the vibration is resolved into discrete bursts, each lasting about 0.5 s and having a complex waveform. The acoustic records from inside the geyser show that these events originate near the top of the water column, where they are shorter (0.1 s or so), have much simpler waveforms, and are recorded slightly earlier than the corresponding seismic signals. Most events consist of two or three pressure pulses, which agree closely with the theoretical sounds from oscillating gas bubbles about 5 cm in radius. Apparently bubbles form at depth, then rise to cooler levels where steam is unstable, whereupon they collapse. The presence of non-condensable gas (probably carbon dioxide) can prevent bubbles from disappearing completely, and cause them to oscillate. The complexity observed

Electron microscopy

Do not adjust your set

This picture may look fuzzy, but that is how the researchers like it. This is the highest resolution image of a crystal structure ever recorded with an electron microscope, resolving columns of silicon atoms 0.78 Å apart. Peter Nellist and Stephen Pennycook (*Phys. Rev. Lett.* 81, 4156–4159; 1998) have achieved this subångström or interatomic imaging by the clever use of incoherent and underfocus techniques.

Unlike light microscopes, where the image resolution approaches the wavelength limit, lens aberrations associated with transmission electron microscopes (TEM) limit the resolution to more than 100 times the wavelength of the electrons used. Conventional TEMs use a coherent beam of electrons to image atoms less than 2 Å apart. Using high-voltage



microscopes to produce faster electrons with shorter wavelengths can improve the imaging up to a point, but the resolution is ultimately restricted by the lens aberrations.

Remarkably, Nellist and Pennycook have been able to achieve subångström

resolution with a microscope operating at a lower voltage and using an incoherent electron source, a method normally used for low-resolution work. They calculate that incoherent electron waves, which are not in phase, interfere and suppress some of the worst effects of the lens aberrations. By combining this technique with underfocusing they have increased the resolving power of the microscope twofold.

The fine structure of crystals cannot easily be probed with methods such as X-ray diffraction, which only resolves regular features of the crystal that contribute to diffraction patterns. With higher-resolution TEM, direct imaging of atomic-scale defects in crystal structures should soon be possible. That should help clarify the picture.

Sarah Tomlin

RICHARD COOMER/PLANET EARTH PICTURES

on seismograms is caused by reverberations at the free surface near the seismometer, and not by source effects.

A surprise in the Old Faithful observations is a lack of evidence of acoustic resonance in the water column. This is the process whereby sounds are magnified by the superposition of echoes reflected many times within the water column, and it has a major role in most models of volcanic tremor. Many analyses, in fact, have bypassed the nonlinear tremor-generation problem entirely, and addressed only linear resonance phenomena in magma bodies, using *ad hoc* mathematical functions as sources. A big problem with resonance models is obtaining impedance contrasts between the fluid and the wall rocks that are high enough for the acoustic waves to be reflected efficiently and create significant resonance. The usual solution has been to hypothesize abundant gas bubbles, which can drastically lower acoustic wave speeds and increase impedance contrasts. But acoustic resonance should be even more important in geysers, where the fluid is water rather than molten rock, and where abundant bubbles can be both seen (on television) and heard. The absence of detectable resonance at Old Faithful casts doubt on its importance in volcanic tremor.

Volcano seismologists have not been devoting all of their attention to geysers, however. They have also deployed digital

seismometers on many volcanoes, and at Mount Semeru (Indonesia)⁶ and Arenal volcano (Costa Rica)⁷ they have now recorded tremor signals strongly indicative of period-doubling cascades⁸, the commonest route to chaos in nonlinear oscillators. Both the waveforms and their spectra show striking regularities, which attracted attention well before their resemblance to the effects of period doubling was recognized. Period doubling is predicted theoretically for flow-induced vibrations³, and these observations suggest that the methods of nonlinear science developed over the past two decades can fruitfully be applied to seismic data from volcanoes. Volcano seismology looks to be poised for rapid progress, thanks to improvements in instrumentation and advances in nonlinear dynamics. □

Bruce R. Julian is in the Earthquake Hazards Team, US Geological Survey, 345 Middlefield Road, MS-977, Menlo Park, California 94025, USA. e-mail: julian@andreas.wr.usgs.gov

1. Kieffer, S. W. *J. Volcanol. Geotherm. Res.* **22**, 59–95 (1984).
2. Kedar, S., Kanamori, H. & Sturtevant, B. *J. Geophys. Res.* **103**, 24283–24299 (1998).
3. Julian, B. R. *J. Geophys. Res.* **99**, 11859–11877 (1994).
4. Chouet, B. A. *Nature* **380**, 309–316 (1996).
5. Hutchinson, R. A., Westphal, J. A. & Kieffer, S. W. *Geology* **25**, 875–878 (1997).
6. Schlindwein, V., Wassermann, J. & Scherbaum, F. *Geophys. Res. Lett.* **22**, 1685–1688 (1995).
7. Hagerty, M. T., Schwartz, S. Y. & Garces, M. A. *J. Volcanol. Geotherm. Res.* (submitted).
8. Julian, B. R. *J. Volcanol. Geotherm. Res.* (submitted).

Neuropharmacology

Cheap date

Caroline J. Small and Stephen R. Bloom

Ever felt that you wanted to change your personality? Well, it may be possible. At last we are getting a handle on many aspects of behaviour — and it's all to do with neuropeptides. On page 366 of this issue, Thiele *et al.*¹ report that the desire to drink alcohol may be inversely related to levels of neuropeptide Y (NPY) in the brain. Thus, a subtle change in brain neuropeptides may dictate a behavioural characteristic.

In humans, sensitivity to ethanol may be genetically influenced^{2,3}. One study has shown that young men with a family history of alcoholism are less susceptible to ethanol than controls, measured subjectively and by looking at body sway, hormonal perturbations and ethanol-related changes in two electrophysiological measures². In a follow-up study a decade later, people with lower sensitivity to alcohol were more likely to be alcoholics³. Ethanol sensitivity in rodent models is also influenced genetically, and strains of mice and rats that differ in their susceptibility to an acute dose of ethanol have been obtained through selective breeding⁴. Knockout mice have allowed

researchers to study these genes further. Mice lacking the 5-hydroxytryptamine (5-HT_{1B}) receptor⁵, the gamma isoform of protein kinase C⁶ or the Fyn tyrosine kinase⁷ have altered responses to ethanol, suggesting that specific genes modulate ethanol-induced behaviour in animals.

Victor Mutt⁸, who died earlier this year and was a pioneer in neuropeptide discovery, was the first to identify NPY, a 36-amino-acid neuropeptide. The potent appetite-stimulating effects of NPY were reported soon after, and it has since been found to facilitate learning and memory, modulate locomotion, inhibit seizures, produce hypothermia, inhibit sexual behaviour, shift circadian rhythms and reduce anxiety (for review see ref. 9). Rats that have been selectively bred for alcohol preference have less NPY in the brain, leading Thiele and colleagues¹ to investigate the ethanol consumption and resistance of mice with altered levels of this neuropeptide.

The authors found that, when allowed to drink freely, NPY knockout mice consumed twice as much ethanol as their wild-type

littermates. Moreover, the knockout animals were resistant to the sedative effects of ethanol. They recovered from an ethanol-induced sleep more rapidly than the wild-type animals, despite having similar plasma-ethanol clearance rates. Thus, mice that lack NPY consume more ethanol and are resistant to its sedative effects.

Thiele *et al.* also examined mice that over-expressed NPY in areas of the brain where this neuropeptide is normally found. These animals showed the reverse — they consumed less ethanol than their wild-type counterparts, and were more sensitive to the ethanol-induced sedation. Previous reports¹⁰ indicate that NPY-deficient mice are anxious, so their increased taste for alcohol could be explained by the relief of such anxiety. However, the authors showed that the mice overexpressing NPY were no different from wild-type controls in measures of anxiety such as fear conditioning and an elevated 'plus maze' test.

Earlier this year Moore *et al.*¹¹ identified a mutant fruitfly with greater ethanol-induced loss of postural control than the wild type. This increased sensitivity to ethanol was independent of both ethanol absorption and metabolism. The mutation, known colourfully as *cheapdate*, was found to occur in an allele of a memory gene, *amnesiac*¹². The *amnesiac* gene is thought to encode a peptide that increases levels of cyclic AMP, and has homology to both the mammalian pituitary adenylate-cyclase-activating peptide and the growth-hormone-releasing hormone. Moore *et al.* suggest that regulation of the cAMP signalling pathway is central to setting ethanol sensitivity in the fruitfly. These interesting findings led Thiele *et al.*¹ to postulate that a physiological function of NPY is to inhibit the production of cAMP in response to alcohol, thereby influencing alcohol consumption.

Antagonists against NPY receptors are currently being proposed as potential anti-obesity agents because they decrease appetite. However, these antagonists may tend to cause an increase in ethanol consumption. Five NPY receptor sub-types have been cloned, and pharmacological studies using fragments and analogues of NPY indicate that there may be more. It is crucial to identify the receptor subtypes that mediate the physiological responses of NPY correctly, to separate the effect on food intake from that on alcohol consumption.

We are now beginning to understand control of behaviour. It is a regulated, physiological function of the brain, like that seen in any other organ. Appetite is, however, a fundamental drive — one animal will kill another in competition for food. And now, NPY, the most potent appetite stimulant yet discovered, has been shown to modulate the desire for alcohol as well. With our increasing understanding of neuropharmacology,

we are on the brink of being able to modify human behaviour. □

Caroline J. Small and Stephen R. Bloom are at the Imperial College of Science, Technology and Medicine, Hammersmith Hospital, London W12 0NN, UK.

e-mail: Sbloom@rpms.ac.uk

1. Thiele, T. E., Marsh, D. J., Ste. Marie, L., Bernstein, I. L. & Palmiter, R. D. *Nature* **396**, 366–369 (1998).
2. Schuckit, M. A. *Alcohol Clin. Exp. Res.* **12**, 465–470 (1988).
3. Schuckit, M. A. *Am. J. Psychiatry* **151**, 184–189 (1994).
4. Crabbe, J. C., Belknap, J. K. & Buck, K. J. *Science* **264**,

- 1715–1723 (1994).
5. Crabbe, J. C. *et al. Nature Genet.* **14**, 98–101 (1996).
6. Harris, R. A. *et al. Proc. Natl Acad. Sci. USA* **92**, 3658–3662 (1995).
7. Miyakawa, T. *et al. Science* **278**, 698–701 (1997).
8. Tatemoto, K., Carlquist, M. & Mutt, V. *Nature* **296**, 659–660 (1982).
9. Turtton, M. D., O'Shea, D. & Bloom, S. R. in *Neuropeptide Y and Drug Development* (eds Grundemar, L. & Bloom, S. R.) 15–39 (Academic, San Diego, 1997).
10. Erickson, J. C., Clegg, K. E. & Palmiter, R. D. *Nature* **381**, 415–421 (1996).
11. Moore, M. S. *et al. Cell* **93**, 997–1007 (1998).
12. Quinn, W. G., Sziber, P. P. & Booker, R. *Nature* **277**, 212–214 (1979).

Ecology

Nutrient consequences of salmon spread

Peter D. Moore

The end of a salmon run is a pretty gloomy affair — mature fish, having expended their final resources of energy on reproduction, wearily donate what remains of their bodies to the recycling processes of the aquatic ecosystem. Ecologists, lacking sentimentality, have asked whether this input of fish biomass to the headwaters of a stream is a significant import of key elements, such as nitrogen and phosphorus, to the area. How far do these fish-based elements move through the aquatic ecosystem, or even beyond? Do the scavengers and predators that feed on dead and dying salmon influence the movements of these elements? These questions have not been easy to answer, but work using stable isotopes in the Pacific Northwest of America, reported by Ben-David, Hanley and Schell¹ in *Oikos*, indicates that the nitrogen from salmon proteins loses none of its migratory properties when released after death.

Ecologists study the nutrient budgets of ecosystems, quantifying the input and output of critical elements, to appreciate how ecosystems function and to develop informed management regimes². For spawning salmon, the nitrogen in their amino acids is largely derived from marine systems where they spend most of their lives. But this nitrogen is delivered to inland wetlands when the fish die. Because each fish carcass contains about 74 g of nitrogen, the total import of this element from salmon migration could be significant. Aquatic invertebrate scavengers and detritivores will find the carcasses a useful resource, as will aquatic microbes, thus enriching the nutrient capital of the wetland itself. But what of surrounding terrestrial and riparian ecosystems? Is there a significant movement of salmon nitrogen into local soils and vegetation? If so, does the nitrogen move, via herbivores, further through the food web of the ecosystem?

To answer these questions, the nitrogen

derived from the fish needs to be traced and distinguished from other sources within the ecosystem. One method that has proved useful in tracing nitrogen sources and pathways is analysis of the stable isotope ¹⁵N in relation to the more abundant ¹⁴N (ref. 3). Only around one in 300 naturally occurring atoms of nitrogen are of the ¹⁵N isotope, but a degree of fractionation (in which one or other of the isotopes becomes differentially concentrated) can occur during physical and biological reactions involving nitrogen. In the case of salmon (marine-based) nitrogen, the $\delta^{15}\text{N}$ (an expression of the ¹⁵N/¹⁴N quotient in relation to a standard) is +13‰, which is richer in ¹⁵N than most terrestrial vegetation (often negative). It should, there-



Figure 1 The nutrients from these sockeye salmon corpses might be about to fertilize the plants in their river's floodplain.

fore, be possible to trace the movement of salmon nitrogen into plants by checking for elevated values of $\delta^{15}\text{N}$.

Ben-David and colleagues¹ have used this approach in their work on Chicagof Island in southeast Alaska. The vegetation there consists of tundra at higher altitudes, and old-growth boreal forest (Sitka spruce *Picea sitchensis* and western hemlock, *Tsuga heterophylla*) at lower altitudes, with blueberries (*Vaccinium* spp.), salmonberry (*Rubus spectabilis*) and devil's club (*Oplopanax horridus*) as the main understorey plants. The authors used 1,000-m transects at right angles to rivers as a sampling system, and analysed the berries and seeds of the understorey plants at intervals along these transects to check the nitrogen-isotope ratios. They found that plants close to the rivers (within about 50 m) tend to have a higher $\delta^{15}\text{N}$ (around +2‰) than those further away (down to −4‰).

Ben-David *et al.* account for these differences by the dispersal of dead salmon over the immediate surrounds of the rivers owing to periodic flooding (Fig. 1). Indeed, observations of carcass density and studies of radio-tagged salmon confirm that a considerable proportion of fish carcasses is carried onto the floodplain by floods. So, the authors conclude that salmon-derived nitrogen enters the vegetation of the riparian ecosystem. When they examined muscle samples from vertebrate herbivores at the riversides (deer mice, voles, shrews and squirrels), they found increased levels of $\delta^{15}\text{N}$ compared with animals from further afield, suggesting that the salmon nitrogen passes along the food chain. Moreover, feeding trials and analyses of the carbon-isotope ratios of these animals indicated that the immediate source of the nitrogen was local herbage rather than the salmon carcasses themselves.

The authors also checked the possible effects of vertebrate predators and scavengers of salmon carcasses — including brown bears, bald eagles, river otters, mink and martens — on the dispersal of fish-based nitrogen. They found a general correlation between evidence for predator activity at a site and the levels of $\delta^{15}\text{N}$ in plants at that site, and concluded that the piscivorous activities of vertebrates could be locally significant in determining nutrient patterns within the ecosystems.

Although this work seems to have established the influence of the salmon run on the nutrient dynamics of this ecosystem, Ben-David and colleagues recognize that other variables need to be examined. It is possible, for example, that the floodplain areas close to rivers have a different pattern of nitrogen cycling. In anaerobic soils, denitrification is likely to be increased, nitrification will be affected, and plant rooting depths will differ. These factors could influence isotope frac-

tionation. Another question, not made clear in this paper, is whether microbial nitrogen fixation — either free-living or symbiotic — occurs in the ecosystem. Two species of alder (*Alnus rubra* and *A. sinuata*) are common in the Pacific Northwest, and these have symbiotic nitrogen-fixing root nodules. They are certainly present in many of the salmon-run catchments of the Cascade Mountains, and, because the microbial fixation of nitrogen produces a $\delta^{15}\text{N}$ of close to zero⁴, this could complicate the picture. We also need more quantitative information on nitrogen inputs (from fixation, precipitation and so on) if the

importance of the salmon run as a fertilizing flush is to be appreciated fully. Only then can we know whether, from the vegetation's point of view, the salmon ran in vain. □

Peter D. Moore is in the Division of Life Sciences, King's College, Campden Hill Road, London W8 7AH, UK.

e-mail: peter.moore@kcl.ac.uk

1. Ben-David, M., Hanley, T. A. & Schell, D. M. *Oikos* **83**, 47–55 (1998).
2. Bormann, F. H. & Likens, G. E. *Science* **155**, 424–429 (1975).
3. Nadelhoffer, K. J. & Fry, B. in *Stable Isotopes in Ecology and Environmental Science* (eds Lathia, K. & Michener, R. H.) 22–44 (Blackwell Science, Oxford, 1994).
4. Kohl, D. H. & Shearer, G. B. *Plant Physiol.* **66**, 51–56 (1980).

Palaeomagnetism

Ancient inclinations

Jean-Pierre Valet and Yves Gallet

In 1839, Carl Friedrich Gauss provided the first mathematical description of the Earth's magnetic field. Since then it has been recognized that the field's main element is an axial dipole aligned with the Earth's rotation axis. That is certainly the case for Earth's later history. But going back to 250 million years ago (Myr) and beyond the evidence becomes much more contrary, as Kent and Smethurst discuss in *Earth and Planetary Science Letters*¹.

A hint of complications comes from the many second-order, multipolar contributions to the field. But it was thought that these contributions (which probably reflect flow patterns at the core–mantle boundary) fluctuate over a few hundred years only, so that the non-axial dipolar field is not expected to retain long-term, persistent features. The axial dipole term dominates², and it has been assumed that, when averaged over enough time, Earth's magnetic field reduces to that of an axial geocentric dipole. The alignment of the dipole field with the rotation axis is a major constraint on the mechanisms that build up the geomagnetic field. Indeed dynamo theory remains preoccupied with fundamental questions of magnetic-field generation and fluid dynamics in the Earth's liquid core. It is also critical for plate tectonic reconstructions, as it provides an ideal frame of reference.

In the past 30 years, palaeomagnetic databases have been developed that offer the possibility of further testing the axial dipole hypothesis. Most analyses^{3–6} have concentrated on the past 5 Myr, because there are abundant data from around the world for this period, and because the motion of tectonic plates can reasonably be ignored. The results confirm that the dipole is indeed a good approximation but surprisingly point to a small (less than 5% of the axial dipole) yet persistent axial quadrupole contribution.

For older periods, things become more

difficult because of loss of resolution caused by uncertainty over site distribution and possible biases introduced by remagnetization. Evans⁷ tackled these problems by taking advantage of the fact that each zonal multipole component has a specific inclination that depends on latitude. So the frequency distribution of palaeomagnetic inclinations can be tested against the theoretical distribution of each zonal field component. The basic idea is that the plate motions have been large enough over periods as long as 100–1,000 Myr to randomize the spatial sampling of the field. Evans averaged the absolute values of the inclinations in $10^\circ \times 10^\circ$ longitude/latitude grids for 11 geological periods to minimize biases in samplings, and observed no significant departure from the distribution predicted by an axial dipole since the end of the Precambrian (550 Myr). Later, Piper and Grant⁸ confirmed this result; but their analysis hinted at some deviation from the axial dipole for older periods.

Kent and Smethurst¹ have now applied the same analysis to the global palaeomagnetic database⁹, the main difference from previous studies being that they have used more data. They find that evidence for the Cenozoic (0–65 Myr) and Mesozoic (65–250 Myr) is still consistent with the axial dipole, but that the Palaeozoic (250–550 Myr) and Precambrian (550–3,500 Myr) exhibit lower inclinations than would be expected from an axial dipolar distribution.

But how good are the data for these older times? Kent and Smethurst have investigated several possibilities that could have generated artefacts. They convincingly show that contamination by younger magnetizations would in fact produce curves in better agreement with an axial dipole. They also argue that the deviation towards low inclinations does not result from inclination 'shallowing' during deposition or compaction of sediments, because for the Precambrian there are

at least double the data from crystalline rocks as from sediments.

It is, however, trickier to identify possible biases caused by the temporal and geographical distribution of the data. The authors agree that the Palaeozoic may have been affected by the tendency for Eurasia and North America to occupy equatorial positions, so validation of the robustness of this analysis is a delicate business. In contrast, rapid continental motions occurred during the Precambrian but in this case the data come almost completely from the Canadian (Laurentia) and Norwegian (Baltica) cratons.

We compared the two distributions obtained from each area separately and found that the results remained much the same. However, there are no more than 200 data points between 2,000 and 3,500 Myr, so it might have been better to restrict analysis to the second half of the Precambrian between 550 and 2,000 Myr. One may wonder whether the results would remain stable over 500-Myr intervals — that is, over successive time periods that would be only twice as long as the Mesozoic. This would also be instructive because most recent reconstructions tend to indicate that Laurentia and Baltica remained in sub-equatorial positions when the Rodinia supercontinent was assembled between about 750 and 1,100 Myr.

In the absence of any other mechanism that can account for their observations, Kent and Smethurst conclude that the anomalous inclination distributions for the Palaeozoic and Precambrian indicate that the time-averaged geomagnetic field was not reduced to a simple axial dipole during these periods. Among several possibilities, they suggest that there were octupole and quadrupole components of the field, as well as the axial dipole. Alternatively, it may be that the geocentric dipole hypothesis still applies but that the anomalous inclinations resulted from migration of the large continental masses towards the Equator in response to perturbations in the Earth's mantle.

The first hypothesis could be related to changes in the boundary conditions of the geodynamo, such as those that stemmed from the growth of the inner core and the onset of compositional convection. It has also been suggested¹⁰ that the formation of the inner core would affect the activity of the geodynamo and induce significant changes in the geomagnetic field strength. From theory¹¹, it seems that the inner core started growing when the temperature at the centre dropped below the freezing point of the core alloy, and the present size of the core means that this process did not begin before 1,700 Myr. One possible objection to a geomagnetic interpretation regarding the inclination distribution in the Palaeozoic is that this would imply a very late formation (or influence) of the inner core. Unfortunately, the

palaeointensity database is still much too incomplete to test further any long-term evolution in the geomagnetic moment.

The second possibility implies the existence of large supercontinents and their migration towards low latitudes in response to geoid anomalies, which could induce tumbling of the Earth with respect to the rotation axis. The authors mention the Rodinia supercontinent as a candidate because it was assembled in the late Precambrian and is supposed to have broken up between 750 and 600 Myr. They also mention Pangaea, which was assembled during the Palaeozoic. It would thus be interesting to detect any relationship between the break-up of these supercontinents and inclination anomalies that would reflect a change in the geoid linked to mass redistribution.

Kent and Smethurst have raised some fascinating issues about the early generation of the geomagnetic field and about the possible influences of the mantle on the field. This is a

case that is far from closed, but tantalizing support for their analysis comes from that small, persistent quadrupole component evident over the past 5 Myr. □

Jean-Pierre Valet and Yves Gallet are at the Institut de Physique du Globe de Paris, 4 Place Jussieu, 75252 Paris Cedex 05, France.

e-mail: valet@ipgp.jussieu.fr

1. Kent, D. V. & Smethurst, M. A. *Earth Planet. Sci. Lett.* **160**, 391–402 (1998).
2. Hulot, G. & LeMouél, J.-L. *Phys. Earth. Planet. Inter.* **82**, 167–183 (1994).
3. Quidelleur, X., Valet, J.-P., Courtillot, V. & Hulot, G. *Geophys. Res. Lett.* **21**, 1639–1642 (1994).
4. Johnson, C. L. & Constable, C. G. *Geophys. J. Int.* **122**, 489–519 (1995).
5. Johnson, C. L. & Constable, C. G. *Geophys. J. Int.* **131**, 643–666 (1997).
6. Carlot, J. & Courtillot, V. *Geophys. J. Int.* **134**, 527–544 (1998).
7. Evans, M. E. *Nature* **262**, 676–677 (1976).
8. Piper, J. D. A. & Grant, S. *Phys. Earth. Planet. Inter.* **55**, 37–53 (1989).
9. McElhinny, M. W. & Lock, J. *Surv. Geophys.* **17**, 557–591 (1996).
10. Stevenson, D. J., Spohn, T. & Schubert, G. *Icarus* **54**, 466–489 (1983).
11. Labrosse, S., Poirier, J.-P. & LeMouél, J.-L. *Phys. Earth Planet. Inter.* **99**, 1–17 (1997).

Muscle

Support for the lever arm

A. F. Huxley

On pages 380 and 383 of this issue, Suzuki *et al.*¹ and Dobbie *et al.*² describe investigations which support current ideas on the way in which changes in the cross-bridges between the myosin and actin filaments of muscle generate force or shortening.

Length changes in muscle take place by relative sliding of two overlapping sets of filaments, composed respectively of myosin and actin. Tension is generated in the overlap regions by cross-bridges formed by the heads of myosin molecules, which attach to an adjacent actin filament, exert force and

detach. Attachment ends when a molecule of ATP binds to the myosin head. This ATP is hydrolysed during the next cycle in steps, some of which are coupled to conformational changes in which the cross-bridge tilts, stretching an elastic component and thus raising the tension.

The only attempt so far at a detailed calculation of the molecular forces between myosin and actin³ suggested that there is a sequence of positions of decreasing potential energy, and that the head as a whole tilts from each position to the next. However, the structure of the myosin head, established by

X-ray diffraction and electron microscopy⁴, suggested that the cross-bridge does not tilt as a whole but that part of the myosin head (called the 'catalytic domain' because it contains a cleft into which ATP or ADP fits) is rigidly attached to the actin filament and that the part which acts as a lever is hinged to this domain. Although that report indicated that the amount of relative movement between the filaments generated by a working stroke would be only about 6 nm, too little to account for the responses of intact muscle fibres to sudden shortening steps⁵, subsequent X-ray analyses of crystals of myosin fragments^{6,7} indicate a stroke of 10–12 nm. Together with about 2 nm contributed by the compliance of the filaments, this would explain the 13 nm taken up during the first few milliseconds after a quick release from isometric contraction⁵.

Suzuki *et al.*¹ describe experiments by fluorescence resonance energy transfer (FRET). Two fluorescent molecules are attached to different points on myosin, and the one working at the shorter wavelength (the donor) is excited by light. The energy of this excited state may transfer to the other (the acceptor), which then emits a photon at its fluorescence wavelength. The extent of this energy transfer varies inversely with the sixth power of the distance between the two fluorophores and so is a sensitive indicator of their separation.

Suzuki *et al.* used a myosin fragment consisting of the catalytic domain together with the presumed hinge and the beginning of the lever. The two fluorophores were attached on either side of the hinge, and the extent of energy transfer was measured without ATP (end of the working stroke). On addition of ATP, which causes the recovery stroke, energy transfer was much reduced. When myosin in the absence of actin binds ATP, it rapidly undergoes an isomerization after which the ATP is hydrolysed to ADP and inorganic phosphate, both of which remain bound until the next working stroke. It follows that the recovery stroke, indicated by the decrease in energy transfer, accompanies or precedes the actual hydrolysis step.

The authors also measured energy transfer in two mutant myosins. With one of these, believed to hold the myosin in the state before the isomerization, the energy transfer was similar to that in rigor; but with the other, the myosin was in the post-isomerization state and the energy transfer was similar to that of the normal myosin with ATP (after the recovery stroke). Therefore the recovery stroke accompanies the isomerization—that is, not hydrolysis itself but the preceding step.

Calculated distances between donor and acceptor agreed with the changes expected from the presumed rotation at the hinge. But the efficiency of fluorescence energy transfer depends on the relative orientations of the fluorophores as well as on their separation.

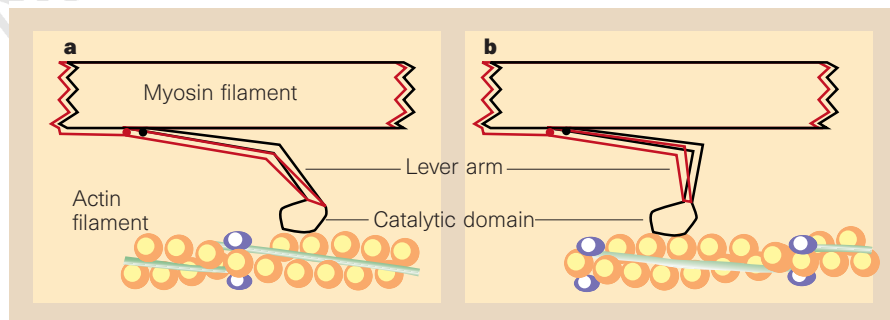


Figure 1 Presumed shape changes in a representative cross-bridge. a, During active contraction at fixed overall length; b, during rigor. Increase of length corresponds to displacement of the myosin filament to the left. The muscle is being subjected to a length oscillation with 3-nm peak-to-peak amplitude of relative sliding of filaments; the red outlines show the situation at the maximum length and the black outlines at the minimum. In a, the centre of mass of the lever arm is to the left of that of the catalytic domain, so that the contour of total mass of the cross-bridge is more spread out in the stretched case (red) than when shortened (black), respectively decreasing and increasing the intensity of the X-ray reflection corresponding to the repeat of myosin molecules along the filament. The intensity therefore fluctuates in opposite phase to the length. In b, the situation is reversed (centre of mass of the lever arm is to the right of that of the catalytic domain) and the intensity fluctuation is in phase with length.

These were unknown and, as is customary, Suzuki *et al.* took an averaged value for their effect on transfer efficiency. Their Fig. 4 (page 382), however, shows that they assume a large change in the relative angular positions of the two fluorophores during the working stroke, so the estimated distances are subject to uncertainty.

Using a variant of FRET, Getz *et al.*⁸ also showed that the working stroke is accompanied by a conformational change. The donor was a molecule containing a terbium atom which is luminescent rather than fluorescent, making a combination that operates over greater separations than ordinary FRET and has other advantages⁹. The fragment of myosin used included the catalytic domain and the whole of the lever. The donor was attached close to the tip of the lever and the fluorescent acceptor to the catalytic domain; actin was present so rigor attachments were formed. On adding ATP (myosin dissociates from actin and recovery stroke occurs), an increase in transfer was found, although sufficient change in separation between donor and acceptor would not be expected from the presumed recovery stroke of the lever. Getz *et al.* concluded that the decrease of transfer was the result of a change in orientation of the acceptor, although other explanations are possible. Their result is a timely warning that FRET results do not unambiguously indicate a change in separation between donor and acceptor.

These observations^{1,8} confirm that a conformational change within the myosin molecule does accompany the working stroke, but do not indicate its extent. The results of Dobbie *et al.*² are complementary: they provide a value for this extent without giving further evidence for a conformational change. An intact fibre from frog muscle was stimulated and a small length oscillation was imposed at a frequency (about 3 kHz) too high for the conformational changes to follow, so that the tension and length changes were in phase. The intensity of the meridional X-ray reflection corresponding to the spacing (14.5 nm) of myosin molecules along their filaments was also recorded and its fluctuations were exactly in opposite phase to the length changes. However, when a similar experiment was done on a fibre in rigor (that is, without ATP so that all cross-bridges were presumably in the state corresponding to the end of the stroke), the X-ray intensity fluctuations were exactly in phase with the length changes.

These results confirm those of Yagi *et al.*¹⁰ at a lower frequency (500 Hz) where phase shifts are appreciable. They imply that the long axis of the cross-bridge swings from one side to the other of the perpendicular to the fibre axis during a working stroke (Fig. 1). This agrees with results from fluorescent probes attached to the lever arm but contradicts a conclusion¹¹ reached (on a scallop

muscle) from electron-paramagnetic-resonance spectra, a method that cannot distinguish between angles on the two sides of the perpendicular.

Dobbie *et al.* explain these X-ray intensity changes quantitatively. In their model (Fig. 3, page 385) the tip of the lever moves by about 9 nm during the working stroke, a little less than the 10–12 nm (suggested by X-ray structures^{6,7}) which matches the amount deduced from sudden length changes (see above). If this 9 nm turns out to be correct it will therefore be necessary to postulate some movement in addition to that from the swing of the lever, perhaps from tilting of the catalytic domain on the actin filament³ or shortening of the connection of the myosin head to the shaft of its filament¹².

All of this is impressive support for the lever-arm hypothesis, but many questions remain. How many of the myosin heads are active at any one time? Is the working stroke a single event or does it go in two or more steps? Answers may come from measurements of the force and amount of movement between a single myosin molecule and an actin filament. But the time resolution of such experiments is inferior to that of experiments on whole fibres, the results are affected by compliance in the actin filaments, and there are big discrepancies between the results from different laboratories. Moreover, experiments are done with myosins from different animals, and at low temperatures where the behaviour of the intact muscle is unknown.

Many will echo Holmes's remark⁶: "The fact that these results fit so nicely with other experiments supporting the lever-arm hypothesis points to their significance". Perhaps so, but results that agree with one's preconceptions need to be scrutinized with especial care. I readily admit that I am inclined to accept results that fit with a working stroke of at least 10 nm suggested long ago by experiments on living muscle⁵, rather than values in the 4–6 nm range that were widely accepted a year ago¹³. □

A. F. Huxley is at Trinity College, Cambridge CB2 1TQ, UK.

1. Suzuki, Y., Yasunaga, T., Ohkura, R., Wakabayashi, T. & Sutoh, K. *Nature* **396**, 380–383 (1998).
2. Dobbie, I. *et al.* *Nature* **396**, 383–387 (1998).
3. Diaz Baños, F. G., Borda, J., Lowy, J. & Svensson, A. *Biophys. J.* **71**, 576–589 (1996).
4. Raymont, I. *et al.* *Science* **261**, 58–65 (1993).
5. Ford, L. E., Huxley, A. F. & Simmons, R. M. *J. Physiol.* **269**, 441–515 (1977).
6. Holmes, K. C. *Curr. Biol.* **7**, R112–R118 (1997).
7. Dominguez, R., Freyzo, Y., Trybus, K. M. & Cohen, C. *Cell* **94**, 559–571 (1998).
8. Getz, E. B., Cooke, R. & Selvin, P. R. *Biophys. J.* **74**, 2451–2458 (1998).
9. Selvin, P. R. & Hearst, J. E. *Proc. Natl Acad. Sci. USA* **91**, 10024–10028 (1994).
10. Yagi, N. *et al.* *J. Synch. Rad.* **3**, 305–312 (1996).
11. Baker, J. E., Brust-Mascher, I., Ramachandran, S., LaConte, L. E. W. & Thomas, D. D. *Proc. Natl Acad. Sci. USA* **95**, 2944–2949 (1998).
12. Harrington, W. F. *Proc. Natl Acad. Sci. USA* **68**, 685–689 (1971).
13. Howard, J. *Nature* **389**, 561–567 (1997).

Daedalus

Fat is a feminist issue

Fat causes a multitude of miseries these days, mainly to women. Emaciation is intensely fashionable. The sub-text of this fashion, says Daedalus, is sexual, but not in the obvious way. Many cultures consider obesity to be sexy; George Orwell commented on the rotundity of the Englishman's secret ideal. This makes good biological sense. Fat is itself a source of oestrogen, so fat women are likely to be sexually enthusiastic — a fact celebrated by a famous English cartoon duo. Conversely, thin women should have a weaker sexual drive. Anorexic women often even cease to menstruate.

So Daedalus reckons that fashionable slimness is really about power. A slim woman hopes to arouse strong desire in the opposite sex, while being immune to it herself — a position of great power. The age-old sex war, augmented by the modern feminist battle, therefore decrees that slimness should be fashionable.

This unstated *realpolitik* of modern fashion leaves fat women in a sorry plight. For not only does fat release sex hormones; sex hormones encourage fat to accumulate, in its usual sites around the female frame. By positive feedback, a fat woman may get ever fatter, more sexually needy, but less fashionably desirable. Meanwhile her slim sisters, sexually switched off by their emaciation, can denounce men for harassing them with unwanted attentions.

Daedalus wants to break this vicious cycle. Many chemicals said to damage animal reproduction (such as DDT) concentrate in body fat. So DREADCO chemists are seeking a drug which dissolves in the fat, but stimulates rather than sabotages its biochemistry. It must boost not only the production of oestrogen (so that even women with little fat will feel sexy) but also leptin — the feedback-control protein which counters obesity by damping production of further fat. Arsenic may give a clue; in low doses it can lead both to emaciation and hypersexuality.

The biochemical development may be long and tortuous. But the final product will splendidly counter the malign influence of modern fashion. With DREADCO's 'Slim and Sexy' fat-activator, fat women will shed their girth without damping their ardour; slim women will no longer sacrifice sexual feeling to fashion; post-menopausal women building their middle-age spread may not need HRT. Militant feminists, of course, will find good reasons to avoid and denounce the product.

David Jones

Graphic gropings

When a flash of inspiration strikes an inventor or a scientist pondering a theory, they scramble for a scrap of paper on which to capture the thought. When these jottings are drawings they become a form of graphic art.

Martin Kemp

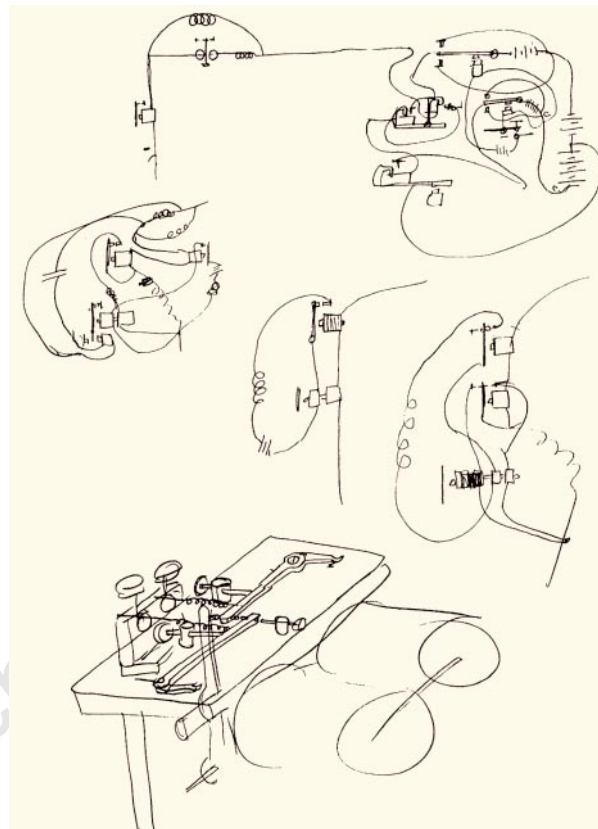
Rough sketches in a notebook, on a blackboard, or even on the back of the proverbial envelope have played central roles in scientific and technological creativity. The rapid act of graphic visualization, the ‘thinking’ doodle, the schematic thought experiment, the improvised diagram, the scribbled solution, are now familiar to us through the publication of scientists’ private papers. These informal fragments seem to bring us closer to the workings of the inventive mind. But it was not always so.

With a few exceptions — Leonardo da Vinci is the most notable example — informal graphic material has not been preserved from earlier eras of science. Some of it was on erasable surfaces, but, more importantly, it was not generally thought worthy of preservation, not least by the originators of the material themselves.

Whereas artists’ sketches had been collectable since the Renaissance, scientists’ notebooks generally attracted little attention until the later part of nineteenth century, and those that did survive were preserved for reasons of archival piety rather than as precious records of inspiration, false starts, changes of mind and so on. The difference speaks of a substantial historical gulf in the attitudes towards creativity in art and science.

It is now broadly accepted that informal jottings can provide fascinating insights into the varied nature of the processes of visualization in different kinds of scientific activity and in individual scientists’ work. The private papers of two of the greatest visualizers, Thomas Edison and Henri Poincaré, show how spontaneous acts of notation on paper can service technological invention no less effectively than they can nourish mathematical theorizing.

Edison consciously projected his own image as a genius of invention — even if his famous “ninety-nine per cent perspiration” proved crucial to his commercial success. His Menlo Park ‘invention factory’ in New Jersey was set up as a crucible for his creative energies. He originated a remarkable string of patented devices there between 1876 and 1883 — an improved telephone, the phonograph, incandescent lighting and so on.



Sketches for a sextuplex set-up in multiple telegraphy, signed by Thomas Edison, James Adams and Charles Batchelor (10 April 1877).

The voluminous notes and drawings signed by Edison and his closest associates display an endless fertility in the conducting of thought experiments. Impulsive jottings of circuits, components, interconnections and new mechanisms, using a variety of graphic conventions, served serendipitous processes of invention, in which ideas could be tested, approved, refined or eliminated.

For his part, Poincaré, uniquely elected to all five sections of the French Academy of Sciences — geometry, mechanics, physics, geography and navigation — became the very model of mathematical creativity. This was exemplified particularly through his

own *Science and Hypothesis* in 1903 and through Edouard Toulouse’s psychological study, published as *Henri Poincaré* in 1910. Poincaré lauded “intuition” as the key to his “ability to perceive the whole argument at a glance” and to determine where the “beauty” of truth lay within nature’s mathematics.

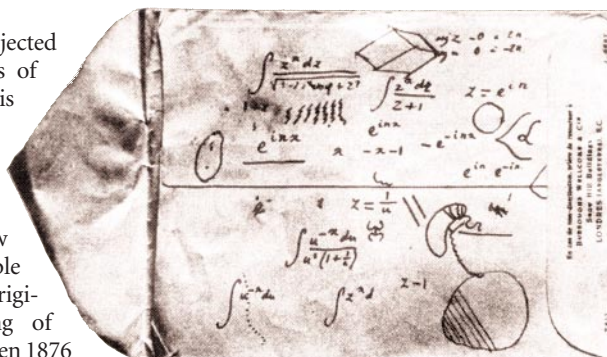
Toulouse describes how, when his subject “searches, he often writes a formula automatically in order to awaken some association of ideas... Poincaré proceeds by sudden blows... During the intervals he assumes... that his unconscious continues the work of reflection”. The mathematician’s manuscripts reflect just this kind of dynamic, associative creativity, darting down formulas and diagrams here and there on any writing surface to hand.

Well might Toulouse express surprise that supremely rational expositions originated from processes that were “seemingly throughout apt for works of pure imagination” in the manner conventionally associated with the arts. □

Martin Kemp is in the Department of the History of Art, University of Oxford, 35 Beaumont Street, Oxford OX1 2PG, UK.

e-mail: martin.kemp@trinity.oxford.ac.uk

Quotations from Poincaré and Toulouse are taken from A. I. Miller, *Insights of Genius* (Copernicus, New York, 1996).



A back-of-the-envelope calculation by Henri Poincaré.

Is beauty in the eye of the beholder?

Why are some humans considered more beautiful than others? Theory suggests that sexually reproducing organisms should choose mates displaying characters indicative of high genotypic or phenotypic quality¹. Attraction to beautiful individuals may therefore be an adaptation for choosing high-quality mates²⁻⁶. Culturally invariant standards of beauty in humans have been taken as evidence favouring such an adaptationist explanation of attraction³⁻⁷; however, if standards of beauty are instead no more than artefacts of culture, they should vary across cultures³⁻⁶. Here we show that male preference for women with a low waist-to-hip ratio (WHR) is not culturally universal, as had previously been assumed.

In the developed world, healthy females have higher levels of oestrogen than testosterone. This causes more fat to be deposited on the buttocks and hips than on the waist⁸, and leads to a low WHR, resulting in a 'waspy-waisted' or gynoid figure (Fig. 1). Females with high WHRs ('apple-shaped' or tubular figures) suffer disproportionately

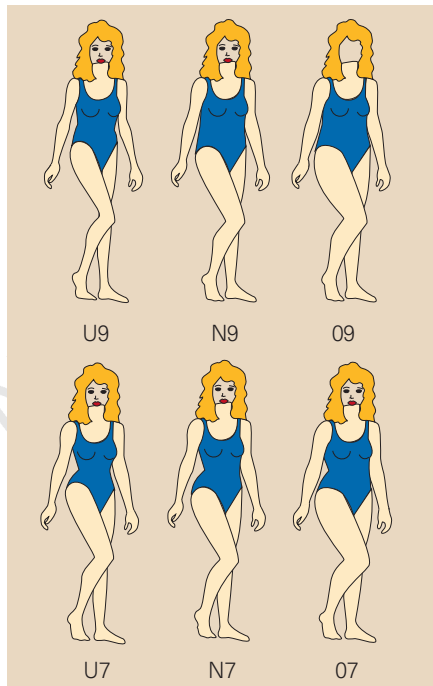


Figure 1 Female figures⁹ depicting two WHR classes (high 0.9 and low 0.7, referred to as 9 and 7) and three weight classes ('overweight', 'normal' and 'underweight', referred to as O, N and U, respectively). A man from Yomybato defined the classes as follows: O9, healthy; O7, had diarrhoea/was skinny in the waist; N9, pallid; N7, had fever, lost weight, especially in the waist; U7, had diarrhoea a few days ago, not skinny but pallid, almost better; U9, pale, almost dead. This figure was adapted with permission from the American Psychological Association (1993).

from a variety of health disorders, including infertility and adult-onset diabetes⁸. Males from several cultures (and times) strongly prefer female figures with a low WHR⁸⁻¹¹, despite otherwise variable preferences for overall weight. The WHR is therefore an accurate indicator of health status and fertility, and male preference for low-WHR females is considered to be one of the best-supported examples of a sexually selected adaptation for assessing mate quality^{2,3,8}.

However, every culture tested so far has been exposed to the potentially confounding influence of western media. Many of the remotest places on Earth have access to television, cinema and advertising posters displaying exceptionally gynoid females draped over desirable products such as cars and beer. Here we test the hypothesis that a population free from such cultural influences would have a different standard of beauty to 'westernized' populations that have been exposed to them.

We assessed WHR preferences in a culturally isolated population of Matsigenka indigenous people from the 1.8-million-hectare Manu Park in southeast Peru. The Matsigenka practise swidden (slash and burn) agriculture, supplemented with fruit and game, using traditional tools. Access to the park's core is restricted to all but scientific and official visitors, and the 300 or so Matsigenka have lived all or most of their lives within the core, so they have not been westernized.

Long-term studies¹² in Yomybato village have familiarized residents with interviews, which were conducted in their native language. Although the people of Yomybato are by no means 'uncontacted', their degree of isolation is about as high as can be obtained today.

WHR preferences of males in Yomybato differed strikingly from those of the US control population (Fig. 2) and from previous results⁸⁻¹¹. Female figures were ranked first by weight and then by WHR. 'Overweight' female O9 (Fig. 1) ranked highest under all three decision criteria: attractiveness, healthiness and preferred spouse. Within weight classes, high-WHR females always ranked significantly higher than low-WHR females (Fig. 2).

To gauge the effect of westernization on WHR preferences, we repeated the WHR test using two more westernized populations: a Matsigenka population living outside Manu Park (Shipetiari), and an even more westernized and ethnically mixed population of Amarakaeri, Huachipaeri and Piro (Alto Madre). Both populations live along the Alto Madre de Dios River, a major

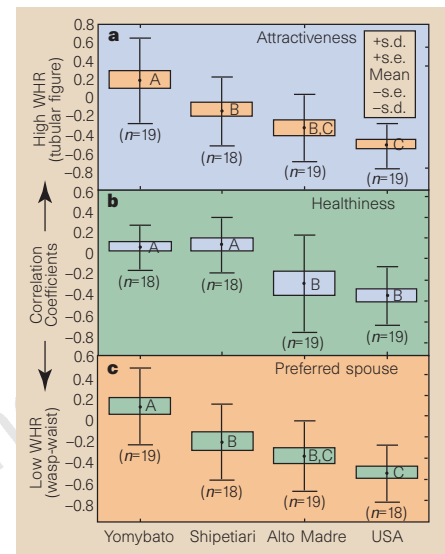


Figure 2 Contrast analyses of WHR preferences. Figures with low and high WHRs were assigned contrast values of +1 and -1, respectively, and, for each informant, correlated against the rank order of the six figures. Correlation coefficients were then pooled within population and decision criteria. Analysis of variance (ANOVA) was used with a post-hoc Tukey HSD for unequal sample sizes, except in the health criterion, in which unequal variances required the use of a Mann-Whitney *U* test and a sequential Bonferroni adjustment of pairwise population comparisons. Populations whose means differed significantly (by $P \leq 0.03$) after table-wide correction are indicated by different letters. In 1996 and 1997, males from several clan lineages were presented with all six figures placed in a circle in random order and asked to rank them from the most to the least beautiful. The process was repeated for healthiness and preferred spouse. Ages in the Yomybato group ranged from 13 to about 60 (median 24.5); Shipetiari, 14 to 50 (26); Alto Madre, 17 to 57 (27); and US, 18 to 60 (24). A fuller description of our results and methods is in preparation.

trading route. The Shipetiari and Yomybato populations are segregates of the same central population that was contacted 20 to 30 years ago.

As in Yomybato, males from Shipetiari grouped female figures by weight and then by WHR, preferred 'overweight' figures, and considered high-WHR females to be healthier (Fig. 2). However, Shipetiari males considered low-WHR females to be more attractive and more desirable as spouses (Fig. 2). Note that Shipetiari males did not consider the healthiest females to be the most attractive or the most desirable spouses, suggesting that WHR preferences may be changing.

Yomybato and Shipetiari women of child-bearing age have high WHRs even

before first pregnancy, and post-childbearing women are thin and have a low WHR.

Alto Madre males did not differ significantly from males from the United States. They grouped female figures by WHR and then by weight: the N7 female (Fig. 1) was generally ranked highest, and low-WHR females were always ranked significantly higher than high-WHR females (Fig. 2). On the gradient of increasing westernization from Yomybato to Alto Madre, WHR and weight preferences become more like those of the US population and previous results^{8–11} (Fig. 2). Our data urgently require replication, because opportunities are disappearing rapidly as more cultures are westernized.

Cultural invariance combined with an adaptive story has been used to support an adaptationist explanation of human beauty (and of other aspects of human psychology¹³). But our results suggest that, when culturally isolated populations are taken into account, some supposedly invariant standards may prove malleable. As a result, many 'cross-cultural' tests in evolutionary psychology may have only reflected the pervasiveness of western media.

A fuller evolutionary theory of human beauty must embrace variation, rather than focusing on 'universal' traits to the exclusion of cultural effects. Of course, physical features are involved in mate choice, and natural selection has probably shaped those preferences. However, in traditional societies, physical features may play a lesser role because mate choice is limited by kinship rules, and potential mates have access to direct information about mate quality, such as age and history of illness, so do not rely on information inferred from physical appearance. In contrast, in industrialized societies, daily exposure to strangers from an early age may increase the importance of using physical features to assess potential mates.

Douglas W. Yu*, Glenn H. Shepard Jr†

*Department of Biology,
Imperial College at Silwood Park,
Ascot, Berkshire SL5 7PY, UK
e-mail: d.yu@ic.ac.uk

†Department of Anthropology,
University of California, Berkeley,
Berkeley, California 94720, USA

- Andersson, M. *Sexual Selection* (Princeton Univ. Press, NJ, 1994).
- Barber, N. *Ethol. Sociobiol.* **16**, 395–424 (1995).
- Symons, D. in *Sexual Nature, Sexual Culture* (eds Abramson, P. R. & Pinkerton, S. D.) 80–120 (Univ. Chicago Press, 1995).
- Buss, D. M. *Behav. Brain Sci.* **12**, 1–49 (1989).
- Cunningham, M. R., Roberts, A. R., Barbee, A. P., Druen, P. B. & Wu, C. H. *J. Pers. Soc. Psychol.* **68**, 261–279 (1995).
- Jones, D. & Hill, K. *Hum. Nature* **4**, 271–296 (1993).
- Perrett, D. I., May, K. A. & Yoshikawa, S. *Nature* **368**, 239–242 (1994).
- Singh, D. *J. Pers. Soc. Psychol.* **65**, 293–307 (1993).
- Singh, D. *Hum. Nature* **6**, 51–68 (1995).
- Furnham, A., Tan, T. & McManus, C. *Pers. Individ. Diff.* **22**, 539–549 (1997).
- Henss, R. *Pers. Individ. Diff.* **19**, 479–488 (1995).
- Shepard, G. H. *J. Amazon. Lang.* **1**, 20–57 (1997).
- Pinker, S. *The Language Instinct* (Harper, New York, 1995).

A developmental model for thalidomide defects

Thalidomide was prescribed as a mild sedative until it was demonstrated in 1961 that it was responsible for inducing congenital malformations in children whose mothers took the drug during pregnancy. Perhaps the most striking aspect of this syndrome is phocomelia (a severe shortening of the limbs). The long bones are shorter than normal and more proximal elements are lost, such that in extreme cases the hand or fingers are attached directly to the shoulder, a condition described as resembling the flipper of a seal^{1,2}. Here I show how these defects can be understood in the context of current models for limb patterning.

Although the pharmacological basis for thalidomide's effect remains controversial, the net result of thalidomide exposure is a decrease in growth of the limb-bud mesenchyme³. From the perspective of pattern formation, the challenge is to understand why the decrease in growth should lead to a specific loss of proximal limb elements. This effect could be viewed as paradoxical, as other perturbations that prevent limb-bud growth, such as removal of the specialized apical ectodermal ridge (AER) at the tip of the limb bud, lead to the opposite patterning defect: deletions of distal structures^{4,5}. These truncations induced by AER removal are explained by the 'progress zone' model of proximodistal patterning⁵, and consideration in this context can also explain the distinct proximodistal defect seen with thalidomide.

According to the progress-zone model, proximodistal structures are specified sequentially under the influence of a continuous signal, now believed to be a fibroblast growth factor (FGF)^{6,7}, produced by the AER. Initially, the entire limb mesenchyme has a proximal identity; left on its own, it would develop into proximal structures.

However, the mesenchyme at the tip of the limb bud under the AER — the progress zone — is exposed to FGF and becomes re-specified to a slightly more distal fate. As limb development proceeds, the progress-zone cells divide and, as a result of this growth, not all of these cells remain within range of the FGF signal. Those too far from the AER maintain their already specified proximal fate, whereas those still in close proximity to the AER are once again re-specified to a still more distal fate (Fig. 1a). This model explains the distal truncations seen following surgical removal of the AER: without the source of the distalizing factor FGF, no further distal patterning occurs and only the proximal elements specified before the AER was removed are formed.

The phocomelia observed following thalidomide treatment can also be understood from this developmental perspective. Thalidomide treatment blocks mesenchymal proliferation in the limb bud³. I propose that this means that the progress zone does not increase in its number of cells, despite continued exposure to FGFs from the AER. As a result, no progress-zone cells end up outside the range of the FGFs, which would be required for them to remember proximally assigned fates. Instead, the entire progress zone is progressively distalized, ultimately programming the entire limb bud to form only the distalmost elements (Fig. 1b).

Consistent with this view, an effect like that of thalidomide can be produced by irradiating the limb mesenchyme, limiting proliferation in the progress zone without affecting the distalizing activity of the AER⁸. Based on these data, a similar mechanism was proposed as a general explanation for phocomelia⁸. Loss of intermediate limb elements, the most common type of limb malformation seen in offspring of women who used thalidomide, is therefore a predicted consequence of altering growth in the progress zone while leaving intact the production of FGFs by the AER.

Many of the pharmacological mechanisms proposed to explain thalidomide's

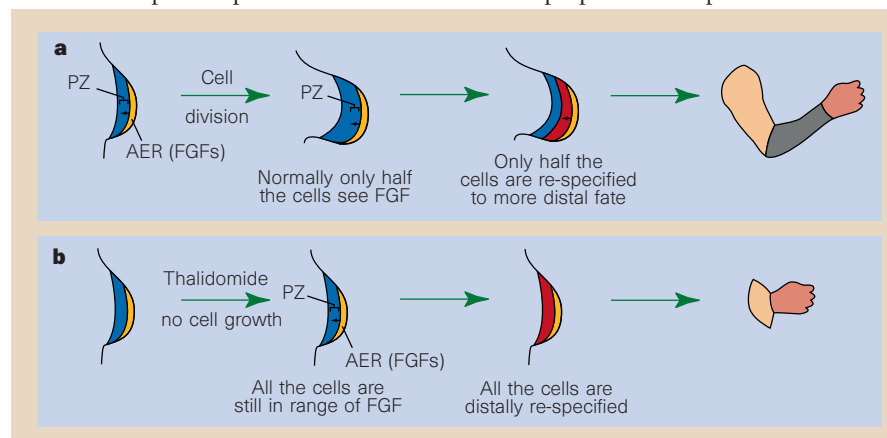


Figure 1 Progress-zone model for thalidomide teratology. **a**, Normal proximodistal pattern specification. **b**, Proximodistal patterning after exposure to thalidomide. PZ, progress zone.

decreasing growth in the early limb bud, such as inhibition of the growth of blood vessels⁵, could also contribute later to decreased growth of the skeletal elements, leading to shortening of the long bones. Although this is not strictly a patterning defect, it is commonly observed in thalidomide cases.

Finally, after the period of exposure to thalidomide, the progress zone would recover, leading to the range of patterning defects occasionally seen, including amelia (absence of limbs), phocomelia, and radial aplasia (lack of development of the radius). The limb patterning defects of thalidomide are therefore the understandable and expected result of decoupling distalization from outgrowth during limb patterning.

Clifford J. Tabin

Department of Genetics,
Harvard Medical School,
200 Longwood Avenue,
Boston, Massachusetts 02115, USA
e-mail: tabin@rascal.med.harvard.edu

1. Lenz, W. *Lancet* **1**, 45 (1962).
2. Toms, D. A. *Lancet* **4**, 400 (1962).
3. Stephens, T. D. *Teratology* **38**, 229–239 (1988).
4. Saunders, J. W. Jr *J. Exp. Zool.* **108**, 363–404 (1948).
5. Summerbell, D. & Lewis, J. H. *J. Embryol. Exp. Morphol.* **33**, 621–643 (1975).
6. Niswander, L., Tickle, C., Vogel, A., Booth, I. & Martin, G. R. *Cell* **75**, 579–587 (1993).
7. Fallon, J. E. *et al. Science* **264**, 104–107 (1994).
8. Wolpert, L., Tickle, C. & Sampford, M. *J. Embryol. Exp. Morphol.* **50**, 175–198 (1979).
9. Kenyon, B. M., Browne, F. & D'Amato, R. J. *Exp. Eye Res.* **64**, 971–978 (1997).

Mistakes not necessary for Müllerian mimicry

In a recent News and Views article¹, Ruxton showed with admirable clarity how the ingenious simulations by MacDougall and Stamp Dawkins² of defensive mimicry in animals highlight the role of cognitive limitations of predators in the generation of relationships involving mimicry. However, Ruxton misrepresents my work³, on which MacDougall and Stamp Dawkins have drawn, thereby giving a misleading impression that mistakes in prey recognition by predators are necessary for the generation of Müllerian mimicry. In fact, the 'virtual predator' on which MacDougall and Stamp Dawkins and others⁴ base their work quite ably generates Müllerian mimicry without depending on predator discrimination errors.

Michael P. Speed

Environmental and Biological Studies,
Liverpool Hope University College,
Hope Park, Liverpool L16 9JD, UK

1. Ruxton, G. D. *Nature* **394**, 833–834 (1998).
2. MacDougall, A. & Stamp Dawkins, M. *Anim. Behav.* **55**, 1281–1288 (1988).
3. Speed, M. P. *Anim. Behav.* **45**, 571–580 (1993).
4. Owen, R. E. & Owen, A. R. G. *J. Theor. Biol.* **109**, 217–247 (1984).

Encapsulated C₆₀ in carbon nanotubes

Pulsed laser vaporization of graphite in the presence of certain metallic catalysts produces both carbon nanotubes and C₆₀ molecules¹. In nanotube production, most of the C₆₀ is removed, along with other residual contaminants, by purification and annealing. It has been suggested that C₆₀ may be trapped inside a nanotube during this elaborate sequence, but this has not been detected.

Here we use high-resolution transmission electron microscopy (HRTEM) to show that closed carbon shells are contained within appropriately sized, single-walled nanotubes. Measurements of the diameter of these endofullerenes suggest that many of them are C₆₀ molecules. Some of them are observed as self-assembled chains with nearly uniform centre-to-centre distances and resemble a nanoscopic peapod. The endofullerenes coalesce into longer capsules under the action of the electron beam.

HRTEM observations were made on purified nanotube material that had been synthesized by pulsed laser vaporization (from A. G. Rinzler and R. E. Smalley). An HRTEM image of an isolated single-walled nanotube is shown in Fig. 1. The image is dark where the electrons encounter the most carbon atoms and are maximally scattered. The parallel lines are the images of opposing tubule walls that are at a tangent, or nearly so, to the electron beam. The lines are separated by between 1.3 and 1.4 nanometres, which is consistent, within the resolution of the microscope, with the diameter of a typical nanotube. Between the lines are ten collinear circles approximately 0.7 nanometres in diameter and spaced 0.3 nanometres from the tubule walls.

These are the expected measurements for C₆₀ contained within the tube and separated from it by a graphitic Van der Waals gap². Furthermore, contrast from the circles is similar to that from the tubule walls, so the constituent atoms are probably carbon

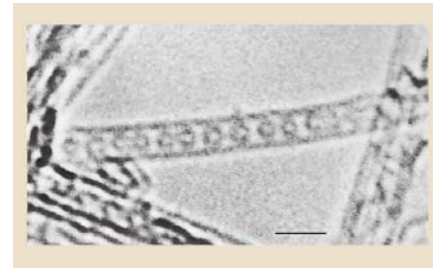


Figure 1 A single-walled carbon nanotube containing a row of closed carbon shells concentric with the tubule axis. The diameter and centre-to-centre spacing of the internal shells are consistent with a chain of C₆₀ molecules. The nanotube is surrounded by a vacuum. Scale bar, 2.0 nanometres.

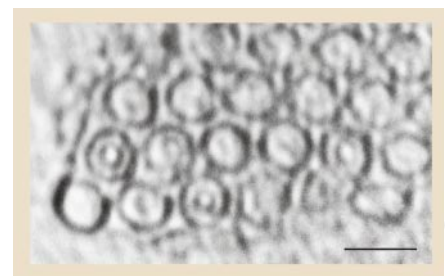


Figure 2 Cross-sectional image of a rope comprising hexagonally packed parallel single-walled carbon nanotubes. At least three tubes contain concentric circular features that are consistent with C₆₀-sized fullerenes. Scale bar, 2.0 nanometres.

rather than a heavier element. Most circles are 1.0 nanometre apart, centre-to-centre, which is the nearest-neighbour spacing in face-centred-cubic C₆₀ (ref. 3), but this separation is slightly less for two pairs.

The HRTEM image shown in Fig. 2 also indicates the presence of contained fullerenes. It shows the cross-section of a rope of single-walled nanotubes that is curved such that the tubule axes are approximately parallel to the electron beam. The image of each tubule is a dark circle 1.3 to 1.4 nanometres in diameter. The tubules are packed hexagonally and are separated by approximately 0.3 nanometres, the expected Van der Waals spacing of tubes in a rope. Three tubes that clearly have an inner concentric circle are interspersed in the lattice. These inner circles cannot be Fresnel artefacts because adjacent tubes in the same focus condition show no internal contrast. The inner circles are about 0.7 nanometres in diameter and are interpreted as the images of contained fullerenes the size of C₆₀ molecules.

After extended exposure to a 100-kilovolt electron beam, the endofullerenes were sometimes seen to coalesce into longer capsules that were oblong, with C₆₀-like caps at the ends. The long parallel walls maintained a nearly uniform spacing of 0.3 nanometres from the walls of the outer tube. Lengths corresponding to the coalescence of three or four spherical molecules were seen. This behaviour may be understood by considering that the energy of binding of a carbon atom to a C₆₀ 'buckyball' is about 0.6 electronvolts less than to a nanotube⁴. For an electron beam of a given energy, this difference means that a carbon atom is more likely to be displaced from C₆₀ than from a nanotube. The coalescence may therefore be the reconfiguration of damaged C₆₀ into more stable capsule structures, with the dimensions constrained by the surrounding nanotube.

Most nanotubes do not contain fullerenes, although many such assemblies were seen. Stable, closed, carbon shells exist inside single-walled carbon nanotubes, and there is strong evidence that many of these

endofullerenes are self-assembled chains of C_{60} . The observation of these structures raises the hope that refined processing techniques can be developed to produce them in large quantities.

Brian W. Smith*, **Marc Monthieux***†, **David E. Luzzi***

*Department of Materials Science and Engineering, University of Pennsylvania, 3231 Walnut Street, Philadelphia, Pennsylvania 19104-6272, USA
e-mail: luzzi@lrm.upenn.edu

†CEMES, UPR A-8011 CNRS, BP 4347, F-31055 Toulouse cedex 4, France

1. Rinzler, A. G. *et al.* *Appl. Phys. A* **67**, 29–37 (1998).
2. Nikolaev, P., Thess, A., Rinzler, A. G., Colbert, D. T. & Smalley, R. E. *Chem. Phys. Lett.* **266**, 422–426 (1997).
3. Heiney, P. A. *J. Phys. Chem. Solids* **53**, 1333–1352 (1992).
4. Yakobson, B. I. & Smalley, R. E. *Am. Sci.* **85**, 324–337 (1997).

How does xenon produce anaesthesia?

Since the discovery that the gas xenon can produce general anaesthesia¹ without causing undesirable side effects, we have remained surprisingly ignorant of the molecular mechanisms underlying this clinical activity of an 'inert' gas. Although most general anaesthetics enhance the activity of inhibitory $GABA_A$ (γ -aminobutyric acid type-A) receptors^{2,3}, we find that the effects of xenon on these receptors are negligible. Instead, xenon potentially inhibits the excitatory NMDA (*N*-methyl-D-aspartate)

receptor channels, which may account for many of xenon's attractive pharmacological properties.

We found that xenon had virtually no effect on $GABA_A$ receptors. Currents activated by 3 μM GABA, both in voltage-clamped cultured rat hippocampal neurons and in voltage-clamped PA3 cells⁴ that stably expressed defined $GABA_A$ subunits, were not significantly affected even by 100% xenon (to function as a human anaesthetic, the half-maximal effective concentration (EC_{50}) is 71% v/v; ref. 5). Xenon also had little effect on functional GABA-releasing synapses in hippocampal neurons, with 80% xenon reducing peak inhibitory postsynaptic currents by only $8 \pm 2\%$. This result indicates that the presynaptic effects of xenon must also be very modest.

Apart from the $GABA_A$ receptor, the only generally accepted neuronal target of conventional anaesthetics is the NMDA receptor. This subtype of glutamate-activated ionotropic channels is implicated in synaptic mechanisms underlying learning, memory and the perception of pain⁶. The NMDA receptor is also believed to be a target of the intravenous general anaesthetic agent ketamine⁷, and possibly nitrous oxide⁸.

We therefore looked at the effects of xenon on NMDA-activated currents in cultured hippocampal neurons. We found that 80% xenon, which will maintain surgical anaesthesia, reduced NMDA-activated currents by about 60% (Fig. 1a), with no significant change in the NMDA EC_{50}

value or Hill coefficient. This non-competitive inhibition indicates that xenon should strongly inhibit neural transmission, despite the high glutamate concentrations in synaptic clefts.

We then tested this in microisland cultures of hippocampal neurons that form synapses with themselves (autapses)⁹. A typical glutamatergic postsynaptic current recorded from a hippocampal neuron is shown in Fig. 1b. The control records show a characteristic biphasic time course, with a fast component mediated by non-NMDA receptors and a much slower component mediated by NMDA receptors. This NMDA receptor-mediated component could be readily identified as it was blocked by the highly selective competitive antagonist AP5 (DL-2-amino-5-phosphonopentanoate)¹⁰.

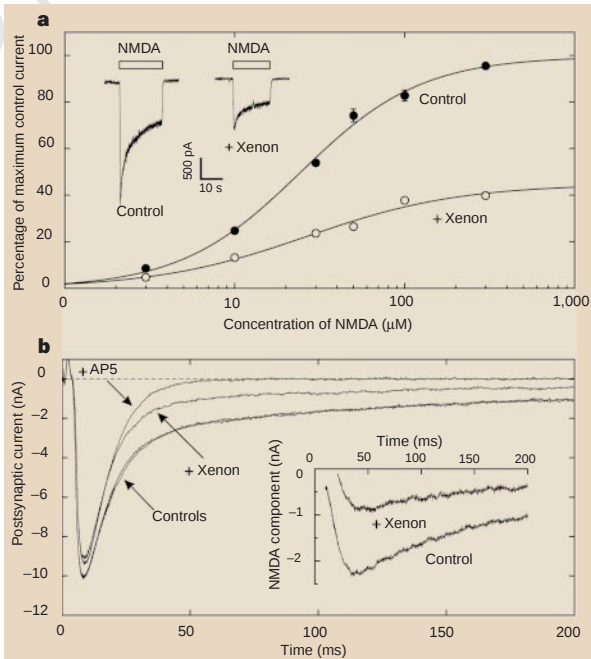
Addition of 200 μM AP5 almost completely blocked the slow component, leaving only a fast component, with a single exponential time course very similar to that of the control fast component. The effect of xenon on the glutamatergic postsynaptic current resembled that of AP5 (Fig. 1b). The slow, NMDA-receptor-mediated component was reduced by over 70%, whereas the fast component barely changed. So, not only did xenon inhibit synaptic NMDA receptors, it had little apparent effect on non-NMDA receptors.

If xenon exerts its effects by inhibiting NMDA receptors, then this explains some important features of its pharmacological profile, particularly as NMDA-receptor antagonists can relieve pain and cause amnesia, which are features of xenon anaesthesia. Like nitrous oxide ('laughing gas'), which may also act, at least partly, on NMDA receptors⁸, xenon can induce a state of euphoria. Other neuronal targets for xenon may emerge, but its powerful inhibition of the NMDA receptor is likely to be instrumental in the anaesthetic and analgesic effects of this 'inert' gas.

N. P. Franks, R. Dickinson, S. L. M. de Sousa, A. C. Hall, W. R. Lieb

Biophysics Section,
The Blackett Laboratory,
Imperial College of Science, Technology and
Medicine, Prince Consort Road,
London SW7 2BZ, UK
e-mail: n.franks@ic.ac.uk

Figure 1 Xenon inhibits NMDA receptors in cultured rat hippocampal neurons. **a**, NMDA activates an inward current (in neurons clamped at -60 mV) with an EC_{50} of 24 ± 2 μM NMDA and a Hill coefficient of 1.2 ± 0.1 . Xenon inhibited the current by approximately 60% but did not significantly change either the EC_{50} or the Hill coefficient. Each data point represents the mean peak current from at least 6 cells. Inset, typical current traces (at 100 μM NMDA) in the presence and absence of xenon. **b**, Xenon selectively inhibits the NMDA-receptor-mediated component of glutamatergic excitatory postsynaptic currents (EPSCs). Neurons were voltage-clamped at -60 mV; synaptic responses were stimulated by a 2-ms depolarizing pulse to $+20$ mV. Control glutamatergic EPSCs displayed a characteristic biphasic decay. The slow component was completely blocked by 200 μM AP5, leaving the fast component almost unaffected. Inset, the NMDA-receptor-mediated component (the difference between the control EPSC and that in the presence of AP5) and its size in the presence of xenon (calculated by taking the difference between the EPSC in the presence of xenon and that in the presence of AP5). Control solutions were equilibrated at room temperature with 80% N_2 and 20% O_2 , and test solutions with 80% Xe and 20% O_2 .



Control glutamatergic EPSCs displayed a characteristic biphasic decay. The slow component was completely blocked by 200 μM AP5, leaving the fast component almost unaffected. Inset, the NMDA-receptor-mediated component (the difference between the control EPSC and that in the presence of AP5) and its size in the presence of xenon (calculated by taking the difference between the EPSC in the presence of xenon and that in the presence of AP5). Control solutions were equilibrated at room temperature with 80% N_2 and 20% O_2 , and test solutions with 80% Xe and 20% O_2 .

1. Cullen, S. C. & Gross, E. G. *Science* **113**, 580–582 (1951).
2. Franks, N. P. & Lieb, W. R. *Nature* **367**, 607–614 (1994).
3. Mihic, S. J. *et al.* *Nature* **389**, 385–389 (1997).
4. Hadingham, K. L. *et al.* *Proc. Natl Acad. Sci. USA* **89**, 6378–6382 (1992).
5. Cullen, S. C., Eger, E. I. II, Cullen, B. F. & Gregory, P. *Anesthesiology* **31**, 305–309 (1969).
6. Rang, H. P., Dale, M. M. & Ritter, J. M. *Pharmacology* 3rd edn (Churchill Livingstone, Edinburgh, 1995).
7. Anis, N. A., Berry, S. C., Burton, N. R. & Lodge, D. Br. *J. Pharmacol.* **79**, 565–575 (1983).
8. Jevtovic-Todorovic, V. *et al.* *Nature Med.* **4**, 460–464 (1998).
9. Bekkers, J. M. & Stevens, C. F. *Proc. Natl Acad. Sci. USA* **88**, 7834–7838 (1991).
10. Watkins, J. C. & Evans, R. H. *Annu. Rev. Pharmacol. Toxicol.* **21**, 165–204 (1981).

When is a dog a DOG?

Similarity and Symbols in Human Thinking

edited by Steven A. Sloman and

Lance J. Rips

MIT Press; 1998. 221 pp. \$25 (pbk)

Jerry Fodor

Cognitive science is a mess. That's largely because the shape a theory of cognition takes depends on the theory of concepts it endorses, and the theory of concepts that (practically all) cognitive scientists endorse is likewise a mess. In this respect, this collection of papers fairly represents the state of the art.

Only two kinds of theories of concepts have ever been proposed, the rationalist and the pragmatist. According to rationalists, to have a concept is to be able to think about the things that it's a concept of. So, to have the concept DOG is to be able to think about dogs. According to pragmatists, to have a concept is to be able to sort things into the ones that the concept applies to and the ones that it doesn't. So, to have the concept DOG is to be able to sort things into dogs and others. These two theories about concepts have many subkinds, each of which has many elaborations, but this rough approximation will do to be getting on with. Well, practically without exception, cognitive scientists are pragmatists about concepts. Here again this book represents the state of the art: all the papers tacitly take for granted that a theory of concepts is a theory of sorting. No hint of an argument for doing so is offered, and no hint of an alternative is implied.

I'm not meaning to suggest that the impression conveyed is of harmonious accord. On the contrary, most of the contributions are polemical, some of them extremely so. Clearly, the authors are vehemently convinced that there is something that they disagree about, and that it matters a lot which of them is right. But so many empirical, ideological and terminological questions are confounded with the main issue that even a sympathetic reader may wonder whether there is and why it does. I aim in what follows to try to help with this. There's an in-house argument among concept pragmatists going on in *Similarity and Symbols in Human Thinking* (S&S) as in cognitive science at large. I'll try to make clear what this argument is about and why I doubt that it can be resolved.

What the argument is about

If the characteristic use of concepts is in sorting, then the characteristic theory of concepts is a model of the sorting process. Pragmatists think of sorting ('categorization') as a matter of matching a concept to a mental representation of the individual whose cate-

gory is to be determined: if they match, then the concept applies to the candidate; if not, then not. Suppose, for example, you're shopping for a tin of gadgets and you wonder whether this tin will do. Here's how to proceed: compare what your gadget concept says about gadgets with what the label on the tin says about what's inside. If what's on the label matches the concept, settle for this tin; if not, try another.

This account is crude, to be sure, but something of the sort is common to all the papers in S&S. Given this shared framework, there are two things left to disagree about: what kinds of mental representations concepts are and how the mind goes about deciding when mental representations match. Unsurprisingly, how one answers either of these questions depends a lot on how one answers the other. Suppose your concept of a gadget is something like a picture of one that you bring along when you go shopping. Then, presumably, sorting is a process of evaluating the 'overall resemblance' between that picture and the one on the label of the tin. (In the jargon of the trade: it's evaluating the 'overall resemblance' between a 'memory image' and a 'perceptual image'.) Or suppose your concept of gadgets is something like a description of them that you've learned. Then sorting is a process of seeing how well the gadget description stored in your memory corresponds to the description on the label. You've learned that gadgets are little square fuzzy things, and the label says that little square fuzzy things are what the can is full of. So, probably, it's a can of gidees.

How close is this to parody? Well, almost nobody these days thinks that concepts could literally be images. For one thing, resemblance couldn't determine extensions in the general case because there are concepts



Clearly, the authors are vehemently convinced that there is something that they disagree about, and that it matters a lot which of them is right.

whose instances aren't things that can be pictured (for example, the concepts NUMBER, WEDNESDAY, QUALM and, for that matter, the concept CONCEPT). Likewise, there are many concepts whose instances can be pictured, but where the sorts of features that a picture can show aren't the ones in virtue of which instances fall under the concept. You can picture a mother, but you can't picture what mothers 'as such' have in common. A picture of Whistler's mother may thus very much resemble Monet's maiden aunt; but Whistler's mother is in the extension of MOTHER and Monet's maiden aunt is not. Analogous problems for the resemblance theory arise about concepts of logical properties such as negation, disjunction, necessity and contingency.

The picture theory of concepts is, in any case, outside the main consensus of the papers in S&S. Almost all the authors assume, more or less explicitly, some version of the theory that concepts and percepts are not images but descriptions. A concept is a list of 'features', stored in memory, that specifies taxonomically relevant properties of the things that the concept applies to. Percepts are likewise feature lists; they represent the properties of a candidate that will be consulted to determine what concept(s) it falls under. Categorizing the candidate consists of evaluating the overlap between these two feature lists. This is all very rough and, I think, also very problematic; I'll return to it presently. Suffice it to say for now that it's common ground in most of S&S. We are finally positioned to see where these papers *don't* agree.

Suppose that a concept specifies the properties in virtue of which things fall into it. There are at least two versions of this sort of view, depending on just what the relation between a thing's having the specified features and its instantiating the concept is taken to be. In particular, the features the concept enumerates might express the properties that all and only the instances of the concept share. Or it might be that they express the properties that all and only the good (that is, the typical) instances of the concept share. On the former view, having a concept is something like knowing its definition; on the latter view, having a concept is something like knowing its prototype.

That is, I think, what practically all the papers in S&S are about: whether concepts are definitions or prototypes. And, I'm afraid, the right answer is that they aren't either; which is why the argument can't be settled. This is a strong claim, to be sure; and I haven't space here to do more than sketch the relevant considerations (for

extended treatments, see my *Concepts*, Oxford University Press, 1998, and chapters 4 and 5 of my *In Critical Condition*, MIT Press, 1998).

Why the argument can't be settled

If a concept is a definition, then there must be a true and necessary biconditional statement that relates the concept to its extension in the way that, for example, the concept BACHELOR is related to the set of (actual and possible) persons who are unmarried and male. Traditional British empiricism broke its heart trying to find examples of such necessary biconditionals. It didn't, and empiricism's current avatars have fared no better.

Over the centuries, the situation has gotten pretty clear. The definition story perhaps works for logico/mathematical concepts (it's necessary that 'P and Q' is true if and only if P is true and Q is true); and it maybe works for a scattering of other cases (necessarily one's mother is one's female parent; necessarily a yawl is a two-masted sailing vessel in which the mizzen is stepped abaft the helm ... and so on.) But for the vast variety of cases, the prospects for such a definition appear quite hopeless. Nobody can define 'cow' or 'concept' or 'friend' or 'table' or 'tree' or 'tenure' or 'basket' or 'foot' or 'since' or 'nice' in this way. There are some half a million words listed in the *Oxford English Dictionary*. At a rough guess, all but a couple of hundred or so are undefinable, at least as far as anybody knows. But if a word hasn't got a definition, then the concept it expresses can't be a definition. Most cognitive scientists now acknowledge that all of this is so. That's why they now adhere to the idea that concepts are prototypes.

But concepts aren't prototypes either. To see why, it helps to approach the issues the rationalist's way, asking 'what do your concepts allow you to think about?' rather than 'what do your concepts allow you to sort?'.

Humans can think about an indefinite variety of things. One can think about: the weather; the weather's being nasty; Sam's being sad that the weather is nasty; it's being obvious that Sam is sad that the weather is nasty; it being surprising that it's obvious that Sam is sad that the weather is nasty; Joan's feeling that it's surprising that it's obvious that Sam is sad that the weather is nasty... and so forth, more or less world without end. One doesn't, in such cases, run out of things that one can think about. In particular, one doesn't run out of concepts to think about them with. This should seem puzzling on reflection. Our minds must, after all, be finite in some sense; but how could a finite mind have an infinite conceptual repertoire?

Here's the standard rationalist answer: although our conceptual repertoires are infinite, they have a finite basis. That is, all but finitely many of our concepts are structurally complex; and the content of a complex con-



Masks of madness

Witches' heads — illustrations from *Beyond Reason: Art and Psychosis. Works from the Prinzhorn Collection* (University of California Press/Hayward Gallery, \$35). These are paintings and drawings by psychiatric patients, assembled between 1918 and 1921 by the German art historian and psychiatrist, Hans Prinzhorn.



cept is determined entirely by its structure together with the content of its constituent concepts. In the most trivial kind of case, for example, the concept BROWN COW is constructed from the constituent concepts BROWN and COW, and this fact entirely determines its identity. That's why the extension of BROWN COW includes all and only things that are in the intersection of the extensions of the concepts BROWN and COW. Systems of representation that work this way are said to be 'compositional'. The fact that our conceptual system is compositional explains not only its productivity, but also an intricate variety of other striking facts about it; that our concepts are 'systematic', for example. (For extended discussion, see Fodor and Pylyshyn, "Connectionism and cognitive architecture: A critical analysis", *Cognition* 28, 3–71; 1988.)

I think it's clear that no theory of concepts has a chance of being true unless it is compatible with concepts being compositional. This is the single most important empirical constraint that theories of concepts are required to meet. By contrast, the issues about compositionality receive no treatment in S&S except for a page in the introduction. It does make *Hamlet* simpler if you leave out the Prince.

The trouble is that prototypes don't com-

pose. That's because the prototype for a concept is a specification of what its good instances have in common; and good-instancehood itself doesn't compose. What's a good instance of RED BARN needn't be a good instance of RED or of BARN. And, even if it is, what's a good instance of RED BARN IN IDAHO needn't be a good instance of RED or of BARN or of RED BARN or of IN IDAHO. What is an instance of a concept depends just on what the concept is. That's why everybody who has the concept RED BARN IN IDAHO knows that its extension includes all and only the (actual and possible) things that are red, barns and in Idaho. But what is a good instance of a concept depends on how things happen to be in the world. In the present case, it depends on how things are with the barns in Idaho; what sorts of barns there happen to be there, and what shades of red they happen to paint them. That's why people who share a concept may fail to share its prototype. And it's likewise why, if you want to know what red barns in Idaho are prototypic, merely reflecting upon the concept won't tell you. You'll just have to go to Idaho and look. This is so even if you happen to know the prototypes of RED, BARN and IN IDAHO.

I want to emphasize how utterly un-

mysterious all this is. From any semantic point of view, the red barns in Idaho are an arbitrary subset of the red barns *tout court*. And there is no particular reason why what is typical of Xs should likewise be typical of an arbitrary subset of Xs. 'Prototype', 'typicality' (and the like) are statistical notions, whereas 'concept', 'extension' and the like are semantic notions. Semantic notions compose but statistical notions don't. So concepts can't be prototypes.

What has gone wrong

I remarked earlier that the debate in S&S is, almost entirely, about whether concepts are definitions or prototypes. Since, for the sorts of reasons I've just been sketching, concepts can't be either definitions or prototypes, it's hard to care much about how the debate in S&S turns out. On the other hand, S&S is a microcosm of current cognitive science views about concepts. The inference has to be that something has gone badly wrong with the current cognitive science views about concepts. What's gone wrong, I think, is the idea that theories of concepts are theories of sorting.

Prima facie, anyhow, being able to sort Xs requires much more than just having the concept X, so how one sorts is a bad test of what concepts one has. I guess I have the concept NASTY WEATHER, and I guess it's intrinsic to that concept that the weather in Patagonia on 12 January 1088 falls under it if and only if the weather there was nasty then. It's because all this is so that I am able to think about the weather in Patagonia at that date; to wonder whether it was nasty; to doubt that it was; to hope that it wasn't... and so forth. This is all just as a rationalist theory of concepts would suppose. But I haven't the foggiest idea how to tell whether the weather in Patagonia on 12 January 1088 was nasty, so I don't know whether to sort it into the NICE WEATHER pile or the NASTY WEATHER pile. Probably nobody knows how to tell; it's a meteorological fact that is lost to us. If you insist on identifying the concept with a sorting procedure — hence having the concept with knowing how to use it to sort — it therefore turns out that I don't have the concept NASTY WEATHER after all; indeed, that nobody does. I do find that implausible.

There is nothing exceptional about this case, and it doesn't hold only for past tense examples. I have the concept of a cloud and the concept of a contrail, but I don't know how to tell which, if either, applies to that little smudgy bit up there. Maybe experts could tell. (But also, maybe not.) If they can, it's not because their concept of a cloud determines the answer and mine doesn't; it's because they know a lot more about clouds and contrails than I do.

Having a concept is one thing, being able to sort with the concept is quite another. Pragmatism about concepts is the practice of

running these two quite different things together, so pragmatism can't be true. There are, however, all sorts of historical and sociological reasons why cognitive scientists are almost all pragmatists about concepts. One is that it's easier to measure what people can sort than it is to measure what they can think about, and measuring things is what most cognitive scientists do for a living. It would thus be both convenient and conducive to full employment if the pragmatist theory of concepts were true. But it's not and the cognitive science it has led to is a mess and it's past time that we kicked the habit. □

Jerry Fodor is in the Department of Philosophy and the Center for Cognitive Science, Rutgers University, New Brunswick, New Jersey 08903, USA.

When maps cost lives

Barrow's Boys

by Fergus Fleming

Granta: 1998. 445 pp. £20

Janet Browne

The recent surge of bookish interest in heroic failures shows that there are any number of incidents in British history ripe for investigation, ranging from the *Sunday Express* reporter who predicted that Franco would never rule Spain to Francis Galton's ludicrous misunderstanding of the Victorian colonial situation which led him to wear hunting pink while exploring deepest Africa. Given the British affection for an underdog, there is a real sense in which failure has underpinned the flow of this country's past just as much as more obvious successes.

Some failures, of course, were genuinely heroic. Fergus Fleming's arresting book about

the British Admiralty takes as its theme the remarkable voyages of exploration that were sent out from Britain during the first half of the nineteenth century. Most of these voyages failed in their purpose, some with disastrous loss of life under appalling conditions. One or two provided a hero's welcome for the returning voyagers, but these were men who were glad merely to be alive. All of them were the brainchild of John Barrow, a solitary, idiosyncratic, irritable figure wedded to his desk in the Admiralty. Barrow had the same sure touch for failure that marked out other, more famous Victorian administrators.

Barrow was second secretary to the navy, a bureaucrat with an unquenchable thirst for discovery, who had joined the Admiralty under Lord Melville, written well-received books on South Africa and China, regularly contributed to the *Quarterly Review*, and survived Melville's fall from power to work under John Wilson Croker for the next 40 years. Much remains unknown about Barrow's motives, and Fleming does well to bring such a shadowy personality to life. Barrow's vision of the world, however, and his evident belief in humanity's power to chart its extremities, can only be guessed at. Yet the nature of his position allowed him to make things happen. Indeed, he made the great age of nineteenth-century exploration happen; and he could rightly be considered as one of the architects of the coming empire.

Peace was Barrow's problem. After the triumphant victory of Trafalgar, it was not at all clear to the naval authorities what they might do with their magnificent taskforce once it was surplus to requirements. Old hands were paid off and turned back onto the streets from which they had first been press-ganged. But the ships and commissioned officers required different treatment. Man or boat, they were expensive to maintain in dock, expensive to scuttle, expensive to keep



To see the world: Porto Praya in the Island of St Jago. From *A Voyage to Conchinchina* by John Barrow.

on half-pay. Even the fighting *Téméraire* was towed up the Thames to be broken up, as depicted by J. M. W. Turner. Barrow's solution was to send them exploring — up the Congo, down the Niger, even to Timbuctoo, and most especially to the far north in search of the fabled north-west passage, the mysterious link between Europe and the Pacific Ocean that would open new trading routes and halve the time taken to reach desirable eastern ports. Barrow's boys were the great naval explorers William Parry, Edward Belcher, John and James Ross, John Franklin, Richard Lander, John Rae and many, many more. One by one, they set out to fulfil Barrow's obsession with filling in blank parts on the Admiralty maps.

The folly of the expeditions described here was extraordinary, the heroism quite stunning. John Ross took a small steamship into Prince Regent Inlet and was stuck in the ice for four winters, spending one season moving no more than 300 yards. Belcher set out with four ships and returned with one. Half of Franklin's first expedition starved to death. George Back went north to save Ross only to discover that, while he iced over at Great Slave Lake, Ross had arrived home. Franklin's third voyage ended with all hands perishing. James Tuckey died in the Congo, Hugh Clapperton died in the Niger, Gordon Laing was murdered in Timbuctoo. Barrow's explorers showed astonishing stoicism and courage in the face of adversity.

Fleming describes some 26 expeditions master-minded by Barrow, including African, Arctic and Antarctic voyages, ranging from Tuckey's inland explorations in 1816 to the final polar voyage in search of Franklin's body, whose death was at last confirmed by Leopold McClintock in 1857. He writes in a thoroughly readable style, full of telling anecdotes, half-amused by the ingenuity of his travellers, critical of their follies and deeply conscious of the extreme rigours they experienced. These stories have been told often enough before, but the strength of Fleming's account is his colourful interlacing of the individual travellers' personal stories. Each voyage spills into the next, with young lieutenants on one expedition ascending the ranks to command another, and with the motif of the solitary, driven, Admiralty man behind them.

What is gained by the personal, however, is partly lost for the domain of science. Few of the scientific achievements of the voyages are described in any detail. It would be hard to unpack the role of the Royal Society of London from these adventurous chapters, for example, in supplying instructions for magnetic and astronomical observations, natural history collections, directives for anthropological encounters and geographical problems to resolve. It would be equally hard to get a clear grasp of the political situation which took Britain from the Napoleonic Wars to William IV (the sailor king) and George

Canning, on to Queen Victoria. Barrow's part in establishing the Royal Geographical Society is scarcely touched on, nor is his baronetcy in 1835. Other voyages of a similarly dramatic nature were sent out by Francis Beaufort of the Hydrographer's Office, most notably the *Beagle*. And more could be said, no doubt, about the exploitative, self-aggrandizing territorial thrust of the British during this period. Science, exploration and national desires were closely intermeshed. The Admiralty's skill in distributing these coiling economic threads across the nineteenth-century globe is indisputable.

But it would be churlish to go on. *Barrow's Boys* is a wonderful story, vividly told. Every amateur yachtsman of my acquaintance will be getting a copy for Christmas. □
Janet Browne is at the Wellcome Institute for the History of Medicine, 183 Euston Road, London NW1 2BE, UK.

To party, or don the anorak?

Mapping Time: The Calendar and Its History

by E. G. Richards
Oxford University Press: 1998. 438 pp. £20, \$33.49

The Calendar: The 5000-Year Struggle To Align the Clock and the Heavens — and What Happened to the Missing Ten Days

by David E. Duncan
Fourth Estate: 1998. 306 pp. £12.99

Jim Bennett

If there is any sensible focus for the millennium, it has to be the calendar, perhaps linked to our interest in what we take to be 'round' numbers. Yet deconstructing the foundation for millennium fever has to go further, for the event occurs in only one particular calendar. Even then, the beginning of the millennium will be identified only according to the current adjustment or finer tuning of that calendar; and all of those doubts come before we might begin to be troubled over which year marks the real beginning.

Has there ever been so much fuss over what is, in the end, a social convention? One reasonable prognostication might be that the vague notion that the millennium will somehow mark the beginning of something more than a different-looking set of numbers is bound to lead to disappointment and disillusion. A harmless response to the hype might be to exchange the party clothes for the anorak, and become surprisingly knowledgeable about the calendar. These books offer alternative ways to begin.

Neither work really admits to a strong millennial mission, although E. G. Richards reassures us that he has noticed something

afloat: "I am aware that soon the world will celebrate the start of a new millennium." But it is hard not to imagine that, at least as far as the publishers are concerned, the big M has influenced these publications, aimed as they are at the popular market.

Mapping Time is reliable in its factual content, but not very analytical, and the facts are dispensed with little sense of historical or cultural context. It reads rather like an encyclopaedia, without adopting that format. Each category is treated in turn in successive chapters without there being a strong narrative structure. It is not a book to be read, but to be consulted on particular topics.

The Calendar, although written in too uncritical a style to be taken seriously as history, has a stronger chronological organization and relies on historical anecdote to keep the narrative moving. Consequently, it is easier to read. It tells the story from the Stone Age to the Gregorian reform of the sixteenth century, and then steps nimbly on to conclude with the atomic clock. It offers a romp through history with calendars as the connecting thread.

Neither book has very seriously scholarly pretensions, although Richards's encyclopaedic approach makes it more strongly weighted with facts and it has a stronger guide to the literature. Both seem to subscribe to a forward, progressive dynamic of science overcoming dogma and superstition with observation and rational thought. This makes their link to the coming millennium somewhat ironic, or perhaps the authors' coyness about any link understandable. □

Jim Bennett is at the Museum of the History of Science, Broad Street, Oxford OX1 3AZ, UK.



Marking time: Boar hunting was a December pastime in the fifteenth century.

The relationship between liquid, supercooled and glassy water

Osamu Mishima & H. Eugene Stanley

That water can exist in two distinct 'glassy' forms—low- and high-density amorphous ice—may provide the key to understanding some of the puzzling characteristics of cold and supercooled water, of which the glassy solids are more-viscous counterparts. Recent experimental and theoretical studies of both liquid and glassy water are now starting to offer the prospect of a coherent picture of the unusual properties of this ubiquitous substance.

At least one 'mysterious' property of liquid water seems to have been recognized 300 years ago¹: whereas most liquids contract as temperature decreases, liquid water begins to expand when its temperature drops below 4 °C. A kitchen experiment demonstrates that the bottom layer of a glass of unstirred iced water remains at 4 °C while colder layers 'float' on top; the temperature at the bottom rises only after all the ice has melted.

The mysterious properties of liquid water become more pronounced in the supercooled region below 0 °C (refs 2–4; Fig. 1). For example, if the coefficient of thermal expansion, isothermal compressibility, and constant-pressure specific heat are extrapolated below the lowest temperatures at which these properties are measurable (about –38 °C at 1 bar) they appear to become infinite at an

unreachable temperature $T_s \approx -45$ °C (refs 2, 5), at which point the entire concept of a 'liquid state' becomes difficult to sustain.

Water is a liquid, but glassy water—also called amorphous ice—can exist when the temperature drops below the glass transition temperature T_g (about 130 K at 1 bar). Although glassy water is a solid, its structure exhibits a disordered liquid-like arrangement. Low-density amorphous ice (LDA) has been known to exist for 60 years (ref. 6), and a second kind of amorphous ice, high-density amorphous ice (HDA), was discovered in 1984^{7–9}. Thus ice has two different amorphous forms. This phenomenon is called polyamorphism, a term meaning that the pure material can exist in more than one amorphous state^{10,11}. All amorphous solid forms of H₂O can be partitioned into these two low-density and high-

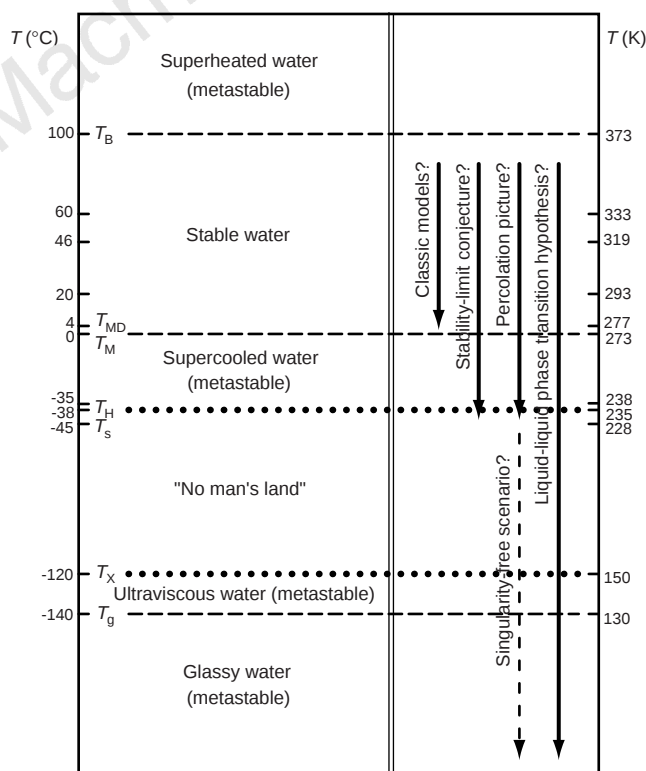


Figure 1 Schematic illustration of different temperature domains, at atmospheric pressure, of H₂O. One domain is stable; the others are metastable. Indicated on the right side of the diagram are the approximate temperature domains which proposed explanations are intended to describe. All the indicated numbers are experimentally observed (and are defined in the text), except the number denoted T_s (about –45 °C) which is a fitting parameter that emerges from assuming the existence of a power-law singularity in measured thermodynamic response

functions. The region between the homogeneous nucleation curve $T_H(P)$ and the crystallization curve $T_X(P)$ is a kind of 'no-man's-land', as experiments on the liquid phase cannot be performed. The temperatures denoted 60 °C, 46 °C, 20 °C and 4 °C indicate the onset of anomalies in the sound velocity, isothermal compressibility, shear viscosity, and density, respectively^{2,3}. T_{MD} denotes the temperature of maximum density, T_B the boiling temperature, T_M the melting temperature, and T_g the glass transition temperature.

density forms, although there may be slight variations in the individual structures depending on the preparation method and procedure^{12–15}.

LDA was originally produced by depositing water vapour onto a cold plate⁶ or by cooling micrometre-sized droplets of liquid water extremely rapidly^{12,16,17}. It was once believed that LDA could be nanocrystalline because of its structural resemblance to the most common form of ice, ice I_h (ref. 18). Now it is agreed that LDA is amorphous, as it transforms to a highly viscous liquid¹⁹ when heated to the known glass transition temperature of about 130 K (refs 20–22).

In contrast, HDA was discovered by compressing ice I_h below a temperature of about 150 K (refs 7, 13, 23). HDA was independently discovered by bombarding an ice- I_h layer with an electron beam^{8,24}. The transformation from a crystalline phase to a (presumably metastable) amorphous phase is called amorphization^{25,26} (Fig. 2). When compressed further at low enough temperatures, HDA crystallizes and becomes a high-density crystalline ice^{23,27}.

In addition to the original preparation methods from the vapour, liquid or crystalline phases of water, LDA can be formed directly from HDA—and vice versa (Fig. 2). This ‘polyamorphic transition’ is accompanied by a surprisingly sharp and surprisingly large volume change—more than 20%—when thermodynamic parameters such as temperature or pressure change infinitesimally. The properties of the amorphous ices change apparently discontinuously at the polyamorphic transition. For example, the shear modulus of LDA decreases on compression, jumps discontinuously at the LDA → HDA transition, and then increases with further compression²⁸. Therefore, the transformation between LDA and HDA appears to be a first-order transition between two different metastable amorphous ‘phases’, rather than a relaxation phenomenon in which an unstable amorphous structure changes continuously from LDA to HDA. Simulations also reveal such polyamorphic transitions^{27,29}.

HDA has a structure similar to that of high-pressure liquid water, suggesting that HDA is a glassy form of high-pressure water^{30–32}. Thus far, however, the glass transition of HDA has not been observed.

There is a continuing discussion about whether pressure-induced amorphization^{13,26,33,34} is conventional ‘two-phase melting’ (crystalline solid becoming liquid) at a temperature below T_g , or is what is called ‘one-phase melting’ (a collapse transition—solid becoming collapsed solid)³⁵ (Fig. 3). Thus it remains unresolved whether one considers HDA to be a glassy state of liquid water or to be a collapsed ‘ill-crystalline’ phase. Regarding HDA as a glass of a high-pressure

liquid, Whalley *et al.*³⁶ predicted from thermodynamic arguments the location of the low-temperature metastable melting line of ice I_h —a prediction later confirmed by experiments¹³—which located the ice I_h melting line above the supposed glass transition temperature T_g . Below T_g , amorphization does not occur on the extrapolation of the two-phase melting line, as once believed³⁷, but rather melting is ‘delayed’ to higher pressures¹³ as suggested in ref. 36 (Fig. 3). This delayed melting can be interpreted as a shift from two-phase melting to one-phase melting.

Here we discuss the status of experimental knowledge of liquid and glassy water in the light of recent theoretical attempts to obtain a unified picture of the liquid, supercooled and glassy states and the transformations between them. These theoretical proposals typically focus on the thermodynamic anomalies exhibited by supercooled water, offering explanations for the singularity-like behaviour that seems to occur at about -45°C . However, it has proved very difficult to find discriminating experimental tests that will allow the theoretical models to be evaluated, and the matter is still unresolved. Computer simulations are unable to resolve the issue either, because water is notoriously hard to simulate realistically and the results can be highly sensitive to the intermolecular potential functions used. Nonetheless, we hope to show that the picture is now becoming clearer, in that some experiments do seem to lend support to specific theoretical models, while simulations have elucidated interrelationships between them. In particular, we can start to discern the origins of water’s anomalous behaviour within the nature of its intermolecular forces. The prospect of a unified picture of water now seems to be in view.

Hypotheses

Many classic ‘explanations’ for the mysterious behaviour of liquid water have been developed^{38–42}, including a simple two-state model dating back to Röntgen⁴³ and a clathrate model dating back to Pauling⁴⁴. Here we shall briefly describe three relatively recent hypotheses under active current discussion.

One of the popular explanations is the stability limit hypothesis⁴⁵, which assumes that the spinodal temperature line $T_s(P)$ in the pressure–temperature (P – T) phase diagram connects at negative P to the locus of the liquid-to-gas spinodal (the limit of metastability) for superheated water (Fig. 4a). It is assumed that liquid water cannot exist when cooled or stretched beyond the line $T_s(P)$.

The data supporting the singularity at T_s are limited to temperatures several degrees above T_s . Hence it is possible that functions have a rapid rise but that there is no singularity, behaviour predicted

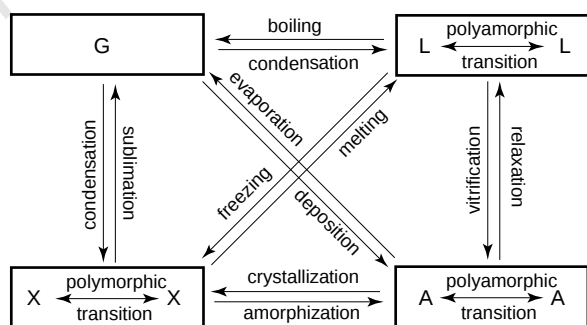


Figure 2 Transitions between gas (G), liquid (L), crystal (X) and amorphous (A) phases of water. We note that the amorphous ‘phase’ is metastable and non-ergodic (meaning a non-equilibrium state in a metastable phase), so the three transitions to and from ‘A’ are different from the other transitions because ‘A’ has many local minima; thus these transitions are not true thermodynamic transitions, but rather depend upon the timescale and other experimental conditions. The transitions to and from ‘A’—and the transitions to and from ‘X’—must be distinguished from those coming from polymorphism and polyamorphism.

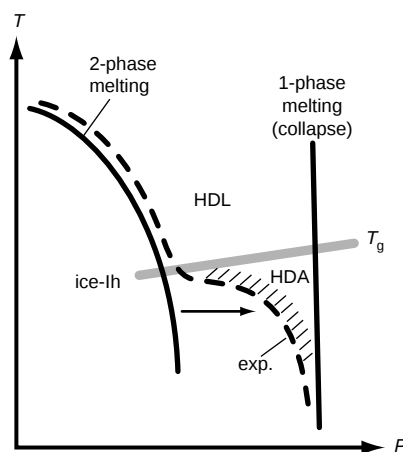


Figure 3 Two-phase melting contrasted with one-phase melting. T – P phase diagram showing the shift of the melting line $T_M(P)$ of ice I_h from ‘two-phase melting’ above T_g to ‘one-phase melting’ below T_g . The experimental data, marked ‘exp.’, follow the dashed line.

by the percolation hypothesis which treats water as a locally structured transient gel comprised of monomers—water molecules—held together by hydrogen bonds whose number increases as temperature decreases⁴⁶. The local ‘patches’ or bonded subdomains⁴⁷ lead to enhanced fluctuations of specific volume and entropy, and to negative cross-correlations of volume and entropy whose anomalies closely match those observed experimentally. The percolation hypothesis has been recognized to be only one example of a more general framework, termed the singularity-free hypothesis (Fig. 4b), that considers the possibility that the observed polyamorphic changes in water are essentially relaxation phenomena that resemble a genuine transition, but are not⁴⁸.

A third recent proposal to rationalize the mysterious properties of water is called the liquid–liquid phase-transition hypothesis (Fig. 4c), which arose from molecular dynamics studies on the structure and equation of state of supercooled water⁴⁹ and has received support from various theoretical approaches^{50–54}. Below the known critical point—with coordinates $T_c = 647$ K, $P_c = 22$ MPa and critical density $\rho_c = 0.328$ g cm⁻³—the high-pressure and high-temperature fluid phase separates into two distinct fluid phases: a low-density gas phase at low pressure and a higher-density liquid phase at high pressure. Below the hypothesized second critical point—with coordinates $T_{c'} \approx 220$ K, $P_{c'} \approx 100$ MPa and $\rho_{c'} \approx 1$ g cm⁻³—the liquid phase separates into two distinct liquid phases: a low-density liquid (LDL) phase at low pressures and a high-density liquid (HDL) at high pressure (Fig. 5a). Water near the known critical point is a fluctuating mixture of molecules whose local structures resemble the liquid and gas phases. Similarly, water near the hypothesized second critical point is a fluctuating mixture of molecules whose local

structures resemble the two phases, LDL and HDL. These enhanced fluctuations influence the properties of liquid water, thereby leading to anomalous behaviour.

Experimental results

Many precise experiments have been performed to test the various hypotheses discussed above, but there is as yet no widespread agreement on which physical picture—if any—is correct. The connection between liquid water and the two amorphous ices predicted by the liquid–liquid phase-transition hypothesis is difficult to prove experimentally because supercooled water freezes spontaneously below the homogeneous nucleation temperature T_H , and amorphous ice crystallizes above the crystallization temperature T_X (refs 55–57). Freezing makes experimentation on the supercooled liquid state between T_H and T_X almost impossible. However, comparing experimental data on amorphous ice at low temperatures with that of liquid water at higher temperatures allows

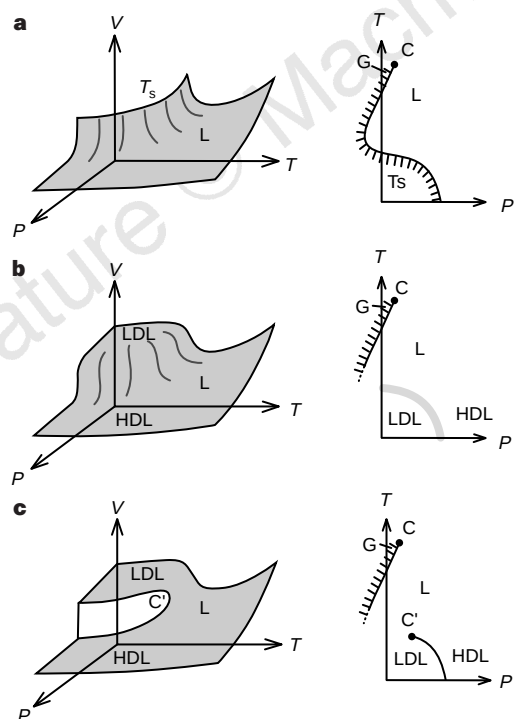


Figure 4 Thermodynamic behaviour of water predicted by three competing theories. The $V(P, T)$ equilibrium equation-of-state surface (left panels) and the projection onto the P – T plane (right panels) for **a**, the stability limit hypothesis, **b**, the percolation picture (and singularity-free hypothesis), and **c**, the liquid–liquid phase-transition hypothesis. Here C denotes the known critical point, C' indicates the hypothesized second critical point, and $T_s(P)$ shows the locus of spinodal temperatures. G and L denote the low-density and high-density fluid phases that exist below C , while LDL and HDL denote the low-density and high-density liquid phases that exist below C' .

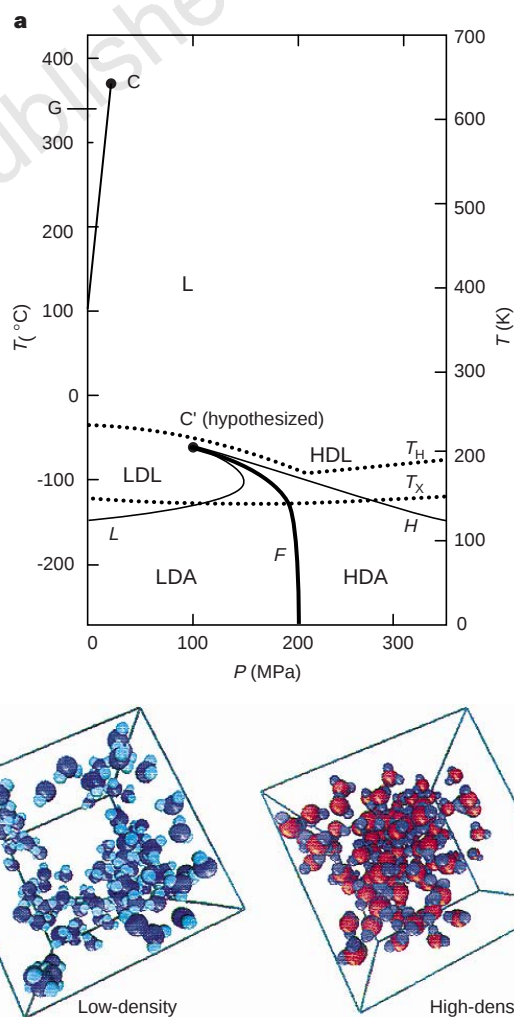


Figure 5 Detailed version of the projection onto the P – T plane of the equilibrium $V = V(P, T)$ surface of Fig. 4c. **a**, The phase relations between liquid water, LDL , HDL , LDA and HDA : C and C' denote the known critical point and the hypothesized ‘second’ critical points respectively, F denotes the line of first-order phase transitions that emanates from C' and separates the high-density and low-density liquid phases that occur for temperatures below $T_{c'}$. The curves denoted L and H are the limits of metastability of the HDA and LDA phases, respectively. **b**, Molecular-dynamics ‘snapshots’ of LDL and HDL , coexisting and separating in liquid water. The subset of water molecules identified in the left panel have a smaller local density than the average, while those shown in the right panel have a larger local density. Courtesy of S. Harrington.

an indirect discussion of the relationship between the liquid and amorphous states. It is found from neutron diffraction studies³⁰ and simulations³² that the structure of liquid water changes towards the LDA structure when the liquid is cooled at low pressures, and changes towards the HDA structure when cooled at high pressures, which is consistent with the liquid–liquid phase-transition hypothesis^{30,32}. The amorphous states (LDA and HDA) are at present considered to be ‘smoothly’ connected thermodynamically to the liquid state if the entropies of the amorphous states are small^{36,58}, and experimental results suggest that their entropies are indeed small⁵⁹.

In principle, it is possible to investigate experimentally the liquid state in the region between T_H and T_X during the extremely short time interval before the liquid freezes to crystalline ice^{13,57,60}. Because high-temperature liquid water becomes LDA without crystallization when it is cooled rapidly at a pressure of 1 bar (refs 12, 16), LDA appears to be directly related to liquid water. A possible connection between liquid water at high pressure and HDA can be seen when ice crystals are melted using pressure¹³. Other experimental results⁵⁷ on the high-pressure ices^{42,61} that might demonstrate a liquid–liquid first-order transition in the region between T_H and T_X have been obtained; these suggest the location of the second critical point to be roughly 100 MPa and 220 K (Fig. 6).

These experimental results are consistent with an LDL–LDA connection at low pressure and an HDL–HDA connection at high pressure as temperature decreases, and clearly show an LDA–HDA transitions at still lower temperatures. They are therefore compatible with both the liquid–liquid first-order transition hypothesis and with the singularity-free hypothesis.

Results from simulations

Many simulation studies of water have been performed, but all are subject to the familiar charge that ‘you get out what you put in’. Water is particularly difficult to simulate because it is a molecular liquid and there is at present no universally agreed precise yet tractable intermolecular potential. Nevertheless, there are some distinct advantages of simulations over experiments. Experiments cannot probe the ‘no man’s land’ that arises from homogeneous nucleation phenomena (Fig. 1), but nucleation does not occur on

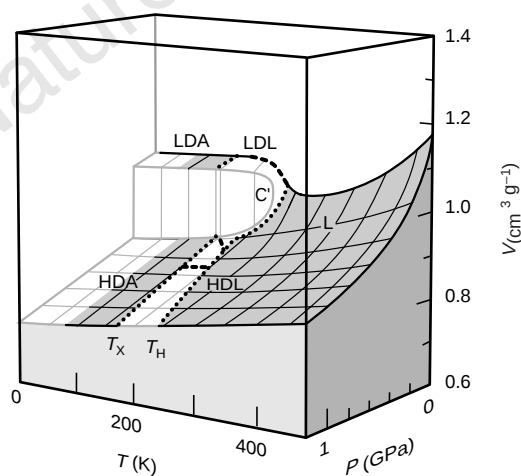


Figure 6 The pressure–volume–temperature relation for H₂O, on the basis of recent experiments⁵⁷. The known LDA and HDA phases are shown, as well as the possible LDL and HDL phases hinted at by simulations⁶⁴ and the interpretation of experimental data⁶⁸; the critical point, above which the LDL and HDL phases become indistinguishable, is also shown. The three dashed lines represent three ‘bridges’ across the ‘no man’s land’ (white) separating the well-studied regions (grey), corresponding to the Mayer LDA experiment at $P = 1$ bar (ref. 12), and the melting lines of ice I_h (ref. 13) and ice XIV (ref. 57). C' denotes the hypothesized LDL–HDL critical point.

the timescale of computer simulations. Hence simulations have the advantage that they can probe the structure and dynamics well below T_H . Of the three hypotheses discussed above, the liquid–liquid phase-transition hypothesis is best supported by simulations, some using the ST2 potential which exaggerates the real properties of water, and others using the SPC/E and TIP4P potentials which underestimate them^{49,62–66}. The precise location of the second critical point is difficult to obtain because the continuation of the first-order line is a locus of maximum compressibility^{62,63,65}.

Further, computer simulations may be used to probe the local structure of water. At low temperatures, many water molecules appear to possess one of two principal local structures, one resembling LDA and the other HDA^{29,49,64,67} (Fig. 5b). Experimental data can also be interpreted in terms of two distinct local structures^{68,69}.

The following three points seem to be consistent both with simulations and with experimental results. (1) Two ‘local’ structures—differing in local density—fluctuate dynamically in high-temperature liquid water, and separate gradually as temperature decreases. (2) At low enough temperatures, the liquid structure resembles more and more the structure of LDA and HDA, depending on the pressure. (3) At extremely low temperatures, LDL and HDL seem to transform, continuously, to LDA and HDA, and one can construct a physically plausible Gibbs-energy surface for water⁵⁷.

Physical arguments

A critical point appears if the pair potential between two particles of the system exhibits a minimum, and Fig. 7a sketches the potential of such an idealized system. At high temperature, the system’s kinetic energy is so large that the potential well does not have an effect, and the system is in a single ‘fluid’ (or gas) phase. At low-enough

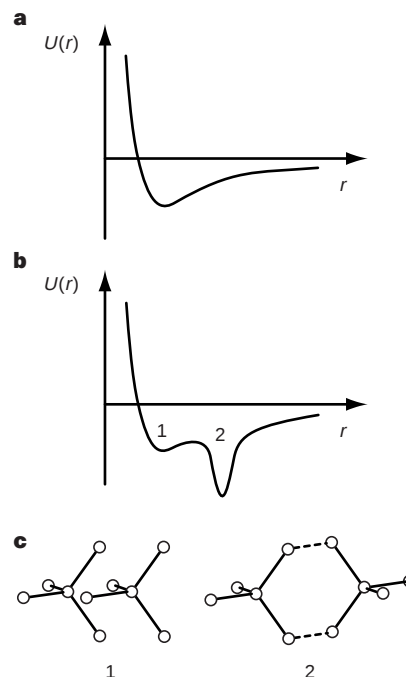


Figure 7 Physical arguments relating to the plausibility of the existence of the known liquid–gas critical point C and the hypothesized LDL–HDL critical point C'. **a**, Idealized system characterized by a pair interaction potential with a single attractive well. **b**, Idealized system characterized by a pair interaction potential whose attractive well has two sub-wells, the outer of which is deeper and narrower. **c**, Two idealized interaction clusters of water molecules in configurations that correspond to the two sub-wells of **b**.

temperature ($T < T_c$) and large-enough pressure ($P > P_c$), the fluid is sufficiently influenced by the minimum in the pair potential that it can condense into the low-specific-volume liquid phase. At lower pressure ($P < P_c$), the system explores the full range of distances—the large-specific-volume gas phase.

Suppose, now, that the potential well has the form shown in Fig. 7b—the attractive potential well shown in Fig. 7a has now bifurcated into a deeper narrow outer sub-well and a more shallow inner sub-well. Such a two-minimum potential can give rise to the occurrence at low temperatures of a second critical point. At high temperature, the system's kinetic energy is so large that the two sub-wells have no appreciable effect of the thermodynamics, and the liquid phase can sample both sub-wells. However, at low enough temperature ($T < T_c$) and not too high a pressure ($P < P_c$), the system must respect the depth of the outer sub-well so the liquid phase 'condenses' into the outer sub-well (the LDL phase). At higher pressure, it is forced into the shallower inner sub-well (the HDL phase). Thus there is a liquid–liquid phase transition at sufficiently low temperatures. It has been appreciated for some time^{70,71} that one can obtain a liquid–liquid phase transition from potentials with generic features similar to the simplified potential shown in Fig. 7b.

The above arguments concern the average or 'thermodynamic' properties, but they may also be useful in anticipating the local properties in the neighbourhood of individual molecules. Consider, again, an idealized fluid with a potential of the form shown in Fig. 7a. Suppose that T is, say, 20% above T_c , so that the macroscopic liquid phase has not yet condensed out. Although the system is not entirely in the liquid state, small clusters of molecules begin to coalesce into the potential well, thereby changing their characteristic interparticle spacing (and hence their local specific volume) and their local entropy, so the fluid system will experience spatial fluctuations characteristic of the liquid phase even though this phase has not yet condensed out of the fluid at $T = 1.2 T_c$. Specific-volume fluctuations are measured by the isothermal compressibility, and entropy fluctuations by the constant-pressure specific heat,

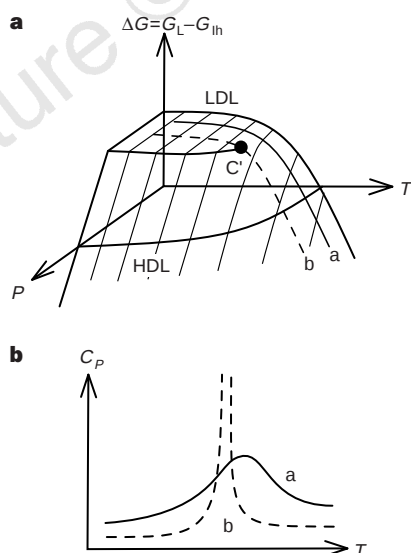


Figure 8 Effect of the hypothesized LDL–HDL critical point C' on two constant-pressure paths, a and b, at pressures just below and just at the critical pressure of C' , respectively. **a**, Schematic of the Gibbs potential surface $G(P, T)$ for stable and metastable water, relative to that of ice I_h , predicted by the liquid–liquid phase-transition hypothesis. **b**, Prediction for the behaviour of the constant-pressure specific heat $C_p = -T(\partial^2 G/\partial T^2)_P$ at 1 atm pressure. As a well-defined line of C_p maxima must emanate from the second critical point, it follows that C_p must display a maximum along each subcritical isobar.

so these two functions should start to increase from the values they would have if there were no potential well at all. As T decreases toward T_c , the magnitude of the fluctuations (and hence of the compressibility and the specific heat) increases monotonically and in fact diverges to infinity as $T \rightarrow T_c$. The cross-fluctuations of specific volume and entropy are proportional to the coefficient of thermal expansion, and this function should start to increase without limit as $T \rightarrow T_c$.

We now consider an idealized fluid with a potential of the form shown in Fig. 7b, and assume that T is below T_c but 20% above T_c , so that the LDL phase has not yet condensed out. The liquid can nonetheless begin to sample the two sub-wells and clusters of molecules will begin to coalesce in each well, with the result that the liquid will experience spatial fluctuations characteristic of the LDL and HDL phase even though the liquid has not yet phase-separated. The specific-volume fluctuations and entropy fluctuations will increase, and so the isothermal compressibility and constant-pressure specific heat will start to diverge. Moreover, if the outer well is narrow, then when a cluster of neighbouring particles samples the outer well it has a larger specific volume and a smaller entropy, so the isothermal expansion coefficient (a measure of the cross-fluctuations of specific volume) is now negative, and approaches $-\infty$ as T decreases.

Now if by chance the value of T_c is lower than the value of T_H , then the actual phase separation discussed above can occur only at temperatures so low that the liquid would have frozen. In this case, the 'hint' of the second critical point is the presence of these local fluctuations whose magnitude would grow as T decreases, but which would never actually diverge if the second critical point is never in fact reached. Functions would be observed experimentally to increase as if they would diverge to ∞ or $-\infty$ but at a temperature below the range of experimental accessibility; we could not directly measure the phase separation itself.

Now consider not the above simplified potential, but rather the potential between water molecules. Presumably the attractive potential well of such a complex nonlinear molecule has structure, and presumably the tetrahedral nature of water dictates that the outermost well corresponds to the ordered configuration and hence has lower entropy. Thus although we do not know the actual form of the intermolecular potential in water, it is not implausible that the same considerations apply as those discussed for the simplified potential of Fig. 7b. Indeed, extensive studies of such pair potentials have been performed, and the existence of the second critical point has been demonstrated in such models^{52,72–74}.

To assess how plausible it is to obtain a bifurcated potential well of the form shown in Fig. 7b, we consider that water can be crudely approximated as a collection of five-molecule groups customarily called Walrafen pentamers (Fig. 7c)⁶⁹. The interaction strengths of two adjacent Walrafen pentamers depends on their relative orientations. The first and the second energy minima of Fig. 7b correspond respectively to configurations 1 and 2 of adjacent Walrafen pentamers shown in Fig. 7c.

Configuration 1 in Fig. 7c is a high-energy, low-specific-volume, high-entropy, non-bonded state; configuration 2 in the same figure is a low-energy, high-specific-volume, low-entropy, bonded state. The difference in local structure between the two energy states depends on whether water molecules enter the empty space between the hydrogen bonds; this is also the difference in the local structure between a high-pressure crystalline ice (such as ice VI or ice VII) and a low-pressure crystalline ice (such as ice I_h)⁴² (Fig. 7c).

The region of the P – T plane along the line continuing from the LDL–HDL coexistence line (when extrapolated to higher temperatures above the second critical point) is the locus of points where the LDL on the low-pressure side and the HDL on the high-pressure side are continuously transforming—it is called the line of compressibility maxima as it is the locus of points where the fluctuations in specific volume are largest. Near this line, two different kinds of

local structures, having either LDL or HDL properties, “coexist”^{76,77}. The entropy fluctuations are also largest near this line and, as illustrated in Fig. 8, the constant-pressure specific heat C_p increases, displaying a lambda-like appearance⁷⁷. Careful measurements and simulations of static and dynamic correlation functions^{75,78–83} may be useful in determining the exact nature of the apparent singular behaviour near 220 K.

Possible implications

Possible forms of polyamorphism in condensed states have been reported, and several examples have been discussed⁸⁴. In addition to H_2O (refs 9, 56, 85), polyamorphism has been studied in Si (ref. 86), and in Al_2O_3 – Y_2O_3 (ref. 87). Liquids and amorphous solids that form network structures (like H_2O , SiO_2 (refs 88, 89), GeO_2 (refs 10, 90), C (refs 91, 92), and other III–V compounds⁹³) are candidates for materials which show a polyamorphic transition and a second critical point. If, indeed, the molten states of SiO_2 and related minerals have polyamorphic transitions under high-pressure and high-temperature conditions inside the Earth, then the transitions may affect the early differentiation and geochemical evolution of this planet⁹⁴.

Which, if any, of the currently discussed hypotheses is correct could affect other fields of research in which water structure and dynamics are important. In chemistry, the structure and properties of aqueous solutions^{95,96} might be viewed differently depending on which model is used. In biology, the correct hypothesis may affect the explanation of the structure and reaction of the liquid water around biological macromolecules⁹⁷. Correct knowledge of the liquid state of water is also important for advancing the art of tissue cryopreservation⁹⁸. Further, in chemical engineering, density and entropy fluctuations near the hypothesized second critical point might be important for aqueous reactions. The structure of water is of particular interest in confined systems⁹⁹, such as the clays that are of geological and biological interest. It has been speculated that polyamorphic transitions in liquid and amorphous solid water may occur inside (or on the surface of) icy bodies in space²⁴, and analogues of polyamorphic transitions have been proposed for the protein folding transition¹⁰⁰.

There are at least three contending hypotheses for rationalizing the puzzling features of liquid water. How to decide between these hypotheses is at a challenge to experimentalists and theoreticians alike, and the puzzling behaviour of liquid water is not yet resolved. Indeed, in the 300 years since the discovery of the density anomaly at 4 °C, many scientific puzzles have been solved, but the mystery of liquid water remains unsolved—despite the widespread importance of this substance. □

O. Mishima is at the National Institute for Research in Inorganic Materials (NIRIM), 1-1, Namiki, Tsukuba, Ibaraki 305-0044, Japan, and at Core Research for Evolutional Science and Technology (CREST), Japan Science and Technology Corporation (JST). H. E. Stanley is at the Center for Polymer Studies and Department of Physics, Boston University, Boston, Massachusetts 02215, USA.

- Waller, R. (trans.) *Essays of Natural Experiments* (original in Italian by the Secretary of the Academie del Cimento); facsimile of the 1684 English translation (Johnson Reprint, New York, 1964).
- Angell, C. A. in *Water: A Comprehensive Treatise* Vol. 7 (ed. Franks, F.) 1–81 (Plenum, New York, 1982).
- Debenedetti, P. G. *Metastable Liquids* (Princeton Univ. Press, 1996).
- Fourkas, J. T., Kivelson, D., Mohanty, U. & Nelson, K. A. (eds) *Supercooled Liquids: Advances and Novel Applications* (ACS Books, Washington DC, 1997).
- Speedy, R. J. & Angell, C. A. Isothermal compressibility of supercooled water and evidence for a thermodynamic singularity at –45°C. *J. Chem. Phys.* **65**, 851–858 (1976).
- Burton, E. F. & Oliver, W. F. The crystal structure of ice at low temperatures. *Proc. R. Soc. Lond. A* **153**, 166–172 (1936).
- Mishima, O., Calvert, L. D. & Whalley, E. ‘Melting’ ice I at 77 K and 10 kbar: a new method of making amorphous solids. *Nature* **310**, 393–395 (1984).
- Heide, H.-G. Observations on ice layers. *Ultramicroscopy* **14**, 271–278 (1984).
- Mishima, O., Calvert, L. D. & Whalley, E. An apparently first-order transition between two amorphous phases of ice induced by pressure. *Nature* **314**, 76–78 (1985).
- Smith, K. H., Shero, E., Chizmeshya, A. & Wolf, G. H. The equation of state of polyamorphic germania glass: A two-domain description of the viscoelastic response. *J. Chem. Phys.* **102**, 6851–6857 (1995).
- Poole, P. H., Grande, T., Sciortino, F., Stanley, H. E. & Angell, C. A. Amorphous polymorphism. *Comput. Mater. Sci.* **4**, 373–382 (1995).

- Mayer, E. in *Hydrogen Bond Networks* (eds Bellissent-Funel, M.-C. & Dore, J. C.) 355–372 (Kluwer Academic, Dordrecht, 1994).
- Mishima, O. Relationship between melting and amorphization of ice. *Nature* **384**, 546–549 (1996).
- Johari, G. P., Hallbrucker, A. & Mayer, E. Two calorimetrically distinct states of liquid water below 150 Kelvin. *Science* **273**, 90–92 (1996).
- Essmann, U. & Geiger, A. Molecular-dynamics simulation of vapor-deposited amorphous ice. *J. Chem. Phys.* **103**, 4678–4692 (1995).
- Brüggeller, P. & Mayer, E. Complete vitrification in pure liquid water and dilute aqueous solutions. *Nature* **288**, 569–571 (1980).
- Dubochet, J. & McDowell, W. A. Vitrification of pure water for electron microscopy. *J. Microsc.* **124**, RP3–RP4 (1981).
- Blackman, M. & Lisgarten, N. D. The cubic and other structural forms of ice at low temperature and pressure. *Proc. R. Soc. Lond. A* **239**, 93–107 (1957).
- Smith, R. S., Huang, C. & Kay, B. D. Evidence for molecular translational diffusion during the crystallization of amorphous solid water. *J. Phys. Chem. B* **101**, 6123–6126 (1997).
- Sugisaki, M., Suga, H. & Seki, S. Calorimetric study of the glassy state. IV. Heat capacities of glassy water and cubic ice. *Bull. Chem. Soc. Jpn* **41**, 2591–2599 (1968).
- McMillan, J. A. & Los, S. C. Vitreous ice: Irreversible transformations during warm-up. *Nature* **206**, 806–807 (1968).
- Johari, G. P., Hallbrucker, A. & Mayer, E. The glass-liquid transition of hyperquenched water. *Nature* **330**, 552–553 (1987).
- Hemley, R. J., Chen, L. C. & Mao, H. K. New transformations between crystalline and amorphous ice. *Nature* **338**, 638–640 (1989).
- Jenniskens, P., Blake, D. F., Wilson, M. A. & Pohorille, A. High-density amorphous ice, the frost on interstellar grains. *Astrophys. J.* **455**, 389–401 (1995).
- Johnson, W. L. Thermodynamic and kinetic aspects of the crystal to glass transformation in metallic materials. *Prog. Mater. Sci.* **30**, 81–134 (1986).
- Sciortino, F. et al. Crystal stability limits at positive and negative pressures, and crystal-to-glass transitions. *Phys. Rev. E* **52**, 6484–6491 (1995).
- Tse, J. S. & Klein, M. L. Pressure-induced phase transformations in ice. *Phys. Rev. Lett.* **58**, 1672–1675 (1987).
- Brazhkin, V. V., Gromnitskaya, E. L., Stal'gorova, O. V. & Lyapin, A. G. Elastic softening of amorphous H_2O network prior to the *hda-lda* transition in amorphous state. *Rev. High Press. Sci. Technol.* **7**, 1129–1131 (1998).
- Poole, P. H., Essmann, U., Sciortino, F. & Stanley, H. E. Phase diagram for amorphous solid water. *Phys. Rev. E* **48**, 4605–4610 (1993).
- Bellissent-Funel, M. C. & Bosio, L. A neutron scattering study of liquid D_2O . *J. Chem. Phys.* **102**, 3727–3735 (1995).
- Bellissent-Funel, M. C., Bosio, L., Hallbrucker, A., Mayer, E. & Sridi-Dorbez, R. X-ray and neutron scattering studies of the structure of hyperquenched glassy water. *J. Chem. Phys.* **97**, 1282–1286 (1992).
- Starr, F. W., Bellissent-Funel, M.-C. & Stanley, H. E. Structural evidence for the continuity of liquid and glassy water. Preprint cond-mat/9811120 at <http://xxx.lanl.gov> (1998); also as *Proc. Les Houches School on Biological Aspects of Water* (ed. Bellissent, M.-C.) (Kluwer, Dordrecht, in the press).
- Sharma, S. M. & Sikka, S. K. Pressure induced amorphization of materials. *Prog. Mater. Sci.* **40**, 1–77 (1996).
- Ponyatovsky, E. G. & Barkalov, O. I. Pressure-induced amorphous phases. *Mater. Sci. Rep.* **8**, 147–191 (1992).
- Tse, J. S. Mechanical instability in ice I_h : A mechanism for pressure-induced amorphization. *J. Chem. Phys.* **96**, 5482–5487 (1992).
- Whalley, E., Klug, D. D. & Handa, Y. P. Entropy of amorphous ice. *Nature* **342**, 782–783 (1989).
- Whalley, E., Mishima, O., Handa, Y. P. & Klug, D. D. Pressure melting below the glass transition: A new way of making amorphous solids. *Proc. NY Acad. Sci.* **484**, 81–95 (1986).
- Bernal, J. D. & Fowler, R. H. A theory of water and ionic solution, with particular reference to hydrogen and hydroxyl ions. *J. Chem. Phys.* **1**, 515–548 (1933).
- Pople, J. A. Molecular association in liquids. II. A theory of the structure of water. *Proc. R. Soc. Lond. A* **205**, 163–178 (1951).
- Frank, H. S. & Wen, W.-Y. Structural aspects of ion-solvent interaction in aqueous solutions: A suggested picture of water structure. *Discuss. Faraday Soc.* **24**, 133–140 (1957).
- Némethy, G. & Scheraga, H. A. Structure of water and hydrophobic bonding in proteins: I. A model for the thermodynamic properties of liquid water. *J. Chem. Phys.* **36**, 3382–3400 (1962).
- Kamb, B. in *Structural Chemistry & Molecular Biology* (eds Rich, A. & Davidson, N.) 507–542 (Freeman, San Francisco, 1968).
- Röntgen, W. C. Ueber die constitution des flüssigen wassers. *Ann. Phys. Chem.* **45**, 91–97 (1892).
- Pauling, L. in *Hydrogen Bonding* (ed. Hadzi, D.) 1–5 (Pergamon, New York, 1959).
- Speedy, R. J. Stability-limit conjecture. An interpretation of the properties of water. *J. Phys. Chem.* **86**, 982–991 (1982).
- Stanley, H. E. & Teixeira, J. Interpretation of the unusual behavior of H_2O and D_2O at low temperatures: Tests of a percolation model. *J. Chem. Phys.* **73**, 3404–3422 (1980).
- Geiger, A. & Stanley, H. E. Low-density patches in the hydrogen-bonded network of liquid water: Evidence from molecular dynamics computer simulations. *Phys. Rev. Lett.* **49**, 1749–1752 (1982).
- Sastry, S., Debenedetti, P., Sciortino, F. & Stanley, H. E. Singularity-free interpretation of the thermodynamics of supercooled water. *Phys. Rev. E* **53**, 6144–6154 (1996).
- Poole, P. H., Sciortino, F., Essmann, U. & Stanley, H. E. Phase behavior of metastable water. *Nature* **360**, 324–328 (1992).
- Ponyatovskii, E. G., Sinitsyn, V. V. & Pozdnyakova, T. A. Second critical point and low-temperature anomalies in the physical properties of water. *JEPT Lett.* **60**, 360–364 (1994).
- Moynihan, C. T. Two species/nonideal solution model for amorphous/amorphous phase transitions. *Mater. Res. Soc. Symp. Proc.* **455**, 411–425 (1997).
- Poole, P. H., Sciortino, F., Grande, T., Stanley, H. E. & Angell, C. A. Effect of hydrogen bonds on the thermodynamic behavior of liquid water. *Phys. Rev. Lett.* **73**, 1632–1635 (1994).
- Borick, S. S., Debenedetti, P. G. & Sastry, S. A lattice model of network-forming fluids with orientation-dependent bonding: equilibrium, stability, and implications from the phase behavior of supercooled water. *J. Phys. Chem.* **99**, 3781–3793 (1995).
- Tejero, C. F. & Baus, M. Liquid polymorphism of simple fluids within a van der Waals theory. *Phys. Rev. E* **57**, 4821–4823 (1998).
- Kanno, H., Speedy, R. & Angell, C. A. Supercooling of water to –92 °C under pressure. *Science* **189**, 880–881 (1975).
- Mishima, O. Reversible first-order transition between two H_2O amorphs at ~0.2 GPa and ~135 K. *J. Chem. Phys.* **100**, 5910–5912 (1994).
- Mishima, O. & Stanley, H. E. Decompression-induced melting of ice IV and the liquid-liquid transition in water. *Nature* **392**, 164–168 (1998).

58. Johari, G. P., Fleissner, G., Hallbrucker, A. & Mayer, E. Thermodynamic continuity between glassy and normal water. *J. Phys. Chem.* **98**, 4719–4725 (1994).
59. Speedy, R. J., Debenedetti, P. G., Smith, R. S., Huang, C. & Kay, B. D. The evaporation rate, free energy, and entropy of amorphous water at 150K. *J. Chem. Phys.* **105**, 240–244 (1996).
60. Bartell, L. S. & Huang, J. Supercooling of water below the anomalous range near 226 K. *J. Phys. Chem.* **98**, 7455–7457 (1994).
61. Bridgman, P. W. The pressure-volume-temperature relations of the liquid, and the phase diagram of heavy water. *J. Chem. Phys.* **3**, 597–605 (1935).
62. Poole, P. H., Sciortino, F., Essmann, U. & Stanley, H. E. The spinodal of liquid water. *Phys. Rev. E* **48**, 3799–3817 (1993).
63. Tanaka, H. Phase behaviors of supercooled water: Reconciling a critical point of amorphous ices with spinodal instability. *J. Chem. Phys.* **105**, 5099–5111 (1996).
64. Harrington, S., Zhang, R., Poole, P. H., Sciortino, F. & Stanley, H. E. Liquid-liquid phase transition: Evidence from simulations. *Phys. Rev. Lett.* **78**, 2409–2412 (1997).
65. Sciortino, F., Poole, P. H., Essmann, U. & Stanley, H. E. Line of compressibility maxima in the phase diagram of supercooled water. *Phys. Rev. E* **55**, 727–737 (1997).
66. Harrington, S., Poole, P. H., Sciortino, F. & Stanley, H. E. Equation of state of supercooled SPC/E water. *J. Chem. Phys.* **107**, 7443–7450 (1997).
67. Shiratani, E. & Sasaki, M. Molecular scale precursor of the liquid-liquid phase transition of water. *J. Chem. Phys.* **108**, 3264–3276 (1998).
68. Bellissent-Funel, M.-C. Is there a liquid-liquid phase transition in supercooled water? *Europhys. Lett.* **42**, 161–166 (1998).
69. Canpolat, M. *et al.* Local structural heterogeneities in liquid water under pressure. *Chem. Phys. Lett.* **294**, 9–12 (1998).
70. Hemmer, P. C. & Stell, G. Fluids with several phase transitions. *Phys. Rev. Lett.* **24**, 1284–1287 (1970).
71. Sadr-Lahijany, M. R., Scala, A., Buldyrev, S. V. & Stanley, H. E. Liquid state anomalies for the Stell-Hemmer core-softened potential. *Phys. Rev. Lett.* **81**, 4895–4898 (1998).
72. Sastry, S., Sciortino, F. & Stanley, H. E. Limits of stability of the liquid phase in a lattice model with water-like properties. *J. Chem. Phys.* **98**, 9863–9872 (1993).
73. Roberts, C. J. & Debenedetti, P. G. Polyamorphism and density anomalies in network-forming fluids: Zeroth- and first-order approximations. *J. Chem. Phys.* **105**, 658–672 (1996).
74. Roberts, C. J., Panagiotopoulos, A. Z. & Debenedetti, P. G. Liquid-liquid immiscibility in pure fluids: Polyamorphism in simulations of a network-forming fluid. *Phys. Rev. Lett.* **77**, 4386–4389 (1996).
75. Shiratani, E. & Sasaki, M. Growth and collapse of structural patterns in the hydrogen bond network in liquid water. *J. Chem. Phys.* **104**, 7671–7680 (1996).
76. Tanaka, H. Fluctuation of local order connectivity of water molecules in two phases of supercooled water. *Phys. Rev. Lett.* **80**, 113–116 (1998).
77. Angell, C. A., Shuppert, J. & Tucker, J. C. Anomalous properties of supercooled water. Heat capacity, expansivity, and proton magnetic resonance chemical shift from 0 to –38°C. *J. Phys. Chem.* **77**, 3092–3099 (1973).
78. Sciortino, F., Fabbian, L., Chen, S.-H. & Tartaglia, P. Supercooled water and the kinetic glass transition: 2. Collective dynamics. *Phys. Rev. E* **56**, 5397–5404 (1997).
79. Sciortino, F., Gallo, P., Tartaglia, P. & Chen, S.-H. Supercooled water and the glass transition. *Phys. Rev. E* **54**, 6331–6343 (1996).
80. Xie, Y., Ludwig, K. F., Morales, G., Hare, D. E. & Sorensen, C. M. Noncritical behavior of density fluctuations in supercooled water. *Phys. Rev. Lett.* **71**, 2051–2053 (1993).
81. Sciortino, F., Poole, P. H., Stanley, H. E. & Havlin, S. Lifetime of the bond network and gel-like anomalies in supercooled water. *Phys. Rev. Lett.* **64**, 1686–1689 (1990).
82. Luzar, A. & Chandler, D. Hydrogen-bond kinetics in liquid water. *Nature* **379**, 55–57 (1996).
83. Starr, F. W., Nielsen, J. & Stanley, H. E. Fast and slow dynamics of hydrogen bonds in liquid water. Preprint cond-mat/9811120 at <http://xxx.lanl.gov> (1998); also as *Comput. Phys. Commun.* (in the press).
84. Brazhkin, V. V., Popova, S. V. & Voloshin, R. N. High-pressure transformations in simple melts. *High Press. Res.* **15**, 267–305 (1997).
85. Mishima, O., Takemura, K. & Aoki, K. Visual observations of the amorphous-amorphous transition in H₂O under pressure. *Science* **254**, 406–408 (1991).
86. Angell, C. A., Borick, S. & Grabow, M. Glass transitions and first order liquid-metal-to-semiconductor transitions in 4-5-6 covalent systems. *J. Non-cryst. Solids* **205**, 463–471 (1996).
87. Aasland, S. & McMillan, P. F. Density-driven liquid-liquid phase separation in the system Al₂O₃-Y₂O₃. *Nature* **369**, 633–636 (1994).
88. Grimsditch, M. Polymorphism in amorphous SiO₂. *Phys. Rev. Lett.* **52**, 2379–2381 (1984).
89. Poole, P. H., Hemmati, M. & Angell, C. A. Comparison of thermodynamic properties of simulated liquid silica and water. *Phys. Rev. Lett.* **79**, 2281–2284 (1997).
90. Itie, J. P. *et al.* Pressure-induced coordination changes in crystalline and vitreous GeO₂. *Phys. Rev. Lett.* **63**, 398–401 (1989).
91. Togaya, M. Pressure dependences of the melting temperature of graphite and the electrical resistivity of liquid carbon. *Phys. Rev. Lett.* **79**, 2474–2477 (1997).
92. Grumbach, M. P. & Martin, R. M. Phase diagram of carbon at high pressures and temperatures. *Phys. Rev. B* **54**, 15730–15741 (1996).
93. Ponyatovsky, E. G. & Pozdnyakova, T. A. The T-P phase diagrams of amorphous GaSb, InSb and InAs. *J. Non-cryst. Solids* **188**, 153–160 (1995).
94. Farber, D. L. & Williams, Q. Pressure-induced coordination changes in alkali-germinate melts: An in situ spectroscopic investigation. *Science* **256**, 1427–1430 (1992).
95. Angell, C. A. & Sare, E. J. Glass-forming composition regions and glass transition temperatures for aqueous electrolyte solutions. *J. Chem. Phys.* **52**, 1058–1068 (1970).
96. Leberman, R. & Soper, A. K. Effect of high salt concentrations on water structure. *Nature* **378**, 364–366 (1989).
97. Wiggins, P. M. Role of water in some biological processes. *Microbiol. Rev.* **54**, 432–449 (1990).
98. Dubochet, J. *et al.* Cryo-electron microscopy of vitrified specimens. *Q. Rev. Biophys.* **21**, 129–228 (1988).
99. Dor, I. In *Correlations and Connectivity* (eds Stanley, H. E. & Ostrowsky, N.) 188–197 (Kluwer Academic, Dordrecht, 1990).
100. Angell, C. A. Landscapes with megabasins: Polyamorphism in liquids and biopolymers and the role of nucleation in folding and folding diseases. *Physica D* **107**, 122–142 (1997).

Acknowledgements. We thank R. J. Hemley, H. Kanno, J. Karbowski, H. Kodama, E. La Nave, R. Sadr-Lahijany, S. Sastry, A. Scala, F. Sciortino, A. Skibinsky, F. W. Starr, Y. Suzuki and M. Yamada for discussions and critical reading of manuscript drafts. This work was supported by the NSF.

Correspondence and requests for materials should be addressed to H.E.S. (e-mail: hes@bu.edu) for theory, and O.M. at NIRIM (e-mail: mishima@nirim.go.jp) for experiment.

Hsp90 as a capacitor for morphological evolution

Suzanne L. Rutherford*† & Susan Lindquist*

*Howard Hughes Medical Institute, University of Chicago, 5841 South Maryland Avenue MC1028, Chicago, Illinois 60637, USA

The heat-shock protein Hsp90 supports diverse but specific signal transducers and lies at the interface of several developmental pathways. We report here that when *Drosophila* Hsp90 is mutant or pharmacologically impaired, phenotypic variation affecting nearly any adult structure is produced, with specific variants depending on the genetic background and occurring both in laboratory strains and in wild populations. Multiple, previously silent, genetic determinants produced these variants and, when enriched by selection, they rapidly became independent of the Hsp90 mutation. Therefore, widespread variation affecting morphogenic pathways exists in nature, but is usually silent; Hsp90 buffers this variation, allowing it to accumulate under neutral conditions. When Hsp90 buffering is compromised, for example by temperature, cryptic variants are expressed and selection can lead to the continued expression of these traits, even when Hsp90 function is restored. This provides a plausible mechanism for promoting evolutionary change in otherwise entrenched developmental processes.

Among the major heat-shock proteins, Hsp90 is unique in its functions¹. It is not required for the maturation or maintenance of most proteins *in vivo*². Most of its many identified cellular targets are signal transducers—cell-cycle and developmental regulators whose conformational instability is relevant to their roles as molecular switches³. Through low-affinity interactions characterized by repeated cycles of binding and release⁴, Hsp90 keeps these unstable signalling proteins poised for activation until they are stabilized by conformational changes associated with signal transduction^{3,5}. Minor changes in amino-acid sequence can have substantial effects on a protein's conformational stability and Hsp90 recognizes structural features common to unstable proteins rather than specific sequence motifs^{2,6,7,37}. Thus, individual members of highly homologous protein families, such as steroid-hormone receptors^{8,9} or cyclin-dependent¹⁰ or Src-family¹¹ kinases, can vary greatly in their dependence on Hsp90.

Studies of yeast illustrate the specificity of Hsp90: at normal temperatures, reductions in Hsp90 levels that have no apparent effects on cell growth or metabolism can completely abolish signalling through Hsp90-dependent pathways^{5,8,11}. Conditions that cause general protein damage can divert Hsp90 from its normal targets to other partially denatured proteins^{2,6,12,13}. Because of its dual involvement with inherently unstable signal transducers on the one hand, and with the cellular response to stress on the other, Hsp90 may link developmental programs to environmental contingency.

Phenotypic variation in Hsp90 mutants

Mutations in the *Drosophila* Hsp90 gene (*Hsp83*) have been independently isolated in three laboratories^{14–16}. The homozygous mutations are lethal^{14,16} and the mutants are maintained as heterozygous stocks. We frequently observed flies with several unusual morphological abnormalities in these stocks (up to 5% in some stocks). Abnormalities also arose when *Hsp83* mutants were outcrossed with standard laboratory strains (Tables 1, 2 and Fig. 1); these occurred in 1–2% of the F₁ flies studied and in a high proportion of the crosses. Of 141 crosses with more than 10 different standard laboratory strains, 53 crosses produced flies with observable defects (174 of 10,400 flies scored, 1.7%). Defects ranged from subtle to severe, involved either one or both sides of the animal, and included body-part transformations, disrupted

abdominal patterning, bristle duplications, deformed eyes or legs and changes in wing shape or venation (Table 1).

The spontaneous appearance of these developmental abnormalities resulted from altered Hsp90 function. First, similar traits arose with mutant *Hsp83* alleles of independent origin (Table 2). Second, when different heterozygous *Hsp83* stocks were crossed together, producing heteroallelic combinations with even lower Hsp90 function (*Hsp83*¹/*Hsp83*²), both the severity and the incidence of some of these abnormal phenotypes were increased (Fig. 1; compare a, c with k, l). Third, when a standard wild-type laboratory strain (*Ore-R*), was raised on food containing a potent, specific inhibitor of Hsp90 (geldanamycin¹⁷) similar abnormalities were produced (Fig. 1q). Of 271 flies raised on the drug, 21 were abnormal

Table 1 Developmental defects associated with Hsp90 deficit

Body part	Code	Description	No. of observations	Temperature (°C)	F ₁ ?	F ₂ ?
Abdomen	A1	Disorganized tergites	14	25	Yes	–
	A2	External trachea?	7	25	No	n.d.
Bristles	B1	Duplications	36	30, 18	Yes	Yes
	B2	Extra scutellar bristles	48	18	Yes	n.d.
	B3	Split scutellars	8	18	Yes	n.d.
	B4	Forked	5	25	No	n.d.
Eyes	E1	Deformed	22	30	Yes	Yes
	E2	Transformed	7	18	No	–
	E3	Smooth	18	18	Yes	Yes
	E4	Rough	16	25	Yes	Yes
	E5	Black facets	24	18	Yes	Yes
	E6	Eyes absent	3	18	Yes	n.d.
Halteres	H1	Ubx transformations	9	25	Yes	Yes
Legs	L1	Deformed	28	18	Yes	Yes
	L2	Transformed	3	18	No	–
Thorax	T1	Disc eversion	12	25	No	–
	T2	Humeral 'balls'	5	26	Yes	n.d.
	T3	Duplication	6	25	No	n.d.
Wings	W1	Small round	26	18	Yes	Yes
	W2	Notched	6	18	Yes	–
	W3	Wing veins	7	18	No	Yes
	W4	Wing border	5	25	No	–
	W5	Transformed	9	18	Yes	n.d.

Developmental abnormalities produced in *Hsp83* mutants, coded according to the part of the fly affected. The approximate number of observations of, and the temperature most frequently producing, each trait are indicated. The observation of at least one cross producing multiple F₁ flies with a given trait is indicated by 'Yes', as is any instance of transmission of the trait to the F₂ generation. A dash indicates not observed; n.d., not done. Many fewer flies were tested at 30 °C, so this situation is under-represented.

† Present address: Developmental Biology Center, University of California at Irvine, 4205 Biological Science II, Irvine, California 92697-2275, USA.

(7.7%); of 443 controls raised on normal food, only one was affected (mildly—it lacked a bristle).

Effect of genetic background

What disrupted normal development in flies with impaired Hsp90 function? We considered three possibilities of increasing interest. First, the mutants might be more sensitive to the environment: as a stress-response factor, wild-type Hsp90 might simply buffer against 'developmental noise' caused by random micro-environmental effects with little or no genetic basis. Second, Hsp90 mutants might exhibit an increased mutation rate: Hsp90 might be directly or indirectly involved in the fidelity of DNA replication. Third,

cryptic genetic variation might be expressed to a greater extent: because it is a chaperone for signal-transduction elements, Hsp90 might normally suppress the expression of genetic variation affecting many developmental pathways.

Further studies supported the third, and most interesting, of these possibilities. Often, when *Hsp83* alleles were crossed to different normal laboratory stocks, several of the F_1 progeny from a given cross shared the same defect, defects distinct from those observed in crosses to other stocks (Table 1 and Fig. 1e–j). Sometimes repeated crosses to the same stocks again produced abnormal flies with the same defects. When affected F_1 flies were crossed together, similarly affected progeny again often appeared in the F_2



Figure 1 Developmental abnormalities associated with Hsp90 deficits. See Table 1 for coding of traits. Deformities appearing in *Hsp83* mutant stocks: **a**, *13F3/TM6B*, deformed fore-leg (code L1) and transformed 2nd leg (L2) with an ectopic sex-comb (arrow); **b**, *P582/TM6B*, deformed eye (E1) with an extra antennae (arrow); **c**, *e1D/TM6B*, smooth eyes (E3) with black facets (E5); **d**, *P582/TM6B*, eye margin transformed into scutellum (E2). Abnormal F_1 hybrids produced from crosses between *Hsp83* mutant stocks and marked laboratory strains: **e**, *e6DX582^{e1}*, left eye has black facets (E5); **f**, *e6AXdpp⁴*, disorganized abdominal tergites (A1); **g**, *e1DXTM3, fiz-lacz*, small wings (W1); **h**, *e3AXIn(2RH)PL,w^m*,

extraneous tissue growing out of tracheal pit (A2, arrow), **i**, *19F2XCdc37^{e4D}*, eyes absent (E6); **j**, *13F3XCdc37^{e1E}*, wing margin material growing into wing; **m**, *19F2X582^{e1}*, deformed eye. Heteroallelic *Hsp83* combination *e1D/9J1*: **k**, severely deformed legs (L1), **l**, severe black-facet phenotype (E5). Abnormal F_1 hybrids produced with wild-type laboratory stocks and *Hsp83* mutants: **n**, *e1D* or *9J1XIR-6*, thickened wing veins (W3); **o**, *P582XSamar-kind*, transformed wing (W5) and extra scutellar bristle (B2, arrow). Abnormalities in wild-type lines raised on geldanamycin: **p**, *IND-6*, notched wings (W2); **q**, *Ore-R*, deformed eye (E1).

Table 2 Strains yielding abnormal F₁ hybrids when crossed with *Drosophila* Hsp90 mutants or treated with geldanamycin

Category	Stock	Description	Hsp83 hybrid/treated with geldanamycin	F ₁ traits	Source	Date
Marked lab. strains	<i>dpp[4]</i>		<i>e3A, e6A, P582</i>	A1, T1	E. Ferguson	4/96
	<i>Hkgsdf</i>		<i>e1D, 9J1</i>	E1	J. Hall	8/89
	<i>ln(1)w[m4], w[m4]</i>		<i>e1D, P582</i>	E1	Mid-America	1/96
	<i>ln(2Rh)PL, w[m]</i>		<i>e1D, e3A, e6A, 9J1</i>	A1, A2, E1, H1, W5	P. Dimitri	3/96
	<i>TM3, ftz-lacZ</i>		<i>e1D</i>	W3	N. Patel	8/94
	<i>w[1118]; 582[e1]</i>	Viable P582 excision	<i>P582, e3A, e1D, e6A, e6D, 19F2</i>	A2, E1, E2, E3, E5, H1, L1, W5	L. Yue	4/95
	<i>w[1118]; 582[e4]</i>	Viable P582 excision	<i>e1D</i>	E5	L. Yue	4/95
	<i>Cdc37[e1C]</i>	T(3;3)	<i>e1D</i>	E5	G. Rubin ¹⁴	9/95
	<i>Cdc37[e1E]</i>	P26(Δ13 amino acids)	<i>19F2, 9J1</i>	T3, W2, W4, W5	G. Rubin ¹⁴	9/95
	<i>Cd37[e4D]</i>	Null; W7(stop codon)	<i>e3A, e4A, e1D, 9F2, 9J1</i>	A1, B2, E6, L1, T2, T3, W1	G. Rubin ¹⁴	9/95
	<i>Cyp-1[317]f</i>	G19D	<i>e1D</i>	E5, W1	S. Rutherford	3/93
WT lab. stocks	<i>CT-1</i>	South Africa, 1954	<i>e1D, 9J1</i>	W1	Mid-America	5/96
	<i>DmAZ</i>	Arizona, 1990	<i>e1D, P582, 582[211], 582[e13]</i>	W5	T. Karr	10/95
	<i>IR-6</i>	Isofemale <i>Ives</i> derivative	<i>e1D, 9J1</i>	E1, E3, W3	G. Gibson ²⁸	6/96
	<i>IS-3</i>	Isofemale <i>Ives</i> derivative	<i>e1D, e6A, 9J1</i>	A1, B1, H1, L1, W5	G. Gibson ²⁸	6/96
	<i>Ore-R</i>	Population cage stock	Geldanamycin	B1, B2, E1, L1, W1, W2	S. Elgin	6/90
	<i>RI-16</i>	Inbred	<i>P582, 9J1</i>	B1, B3, L1	T. Mackay	8/97
	<i>RI-20</i>	Inbred	<i>9J1</i>	L1	T. Mackay	8/97
	<i>RI-25</i>	Inbred	<i>9J1</i>	T3	T. Mackay	8/97
	<i>RI-27</i>	Inbred	<i>P582, 9J1</i>	T2	T. Mackay	8/97
	<i>Samarkind</i>	Inbred	<i>P582</i>	L1, T1, W5	T. Mackay	8/97
	<i>wol</i>	<i>Ore-R</i> derivative	<i>P582</i>	H1	T. Karr	10/95
Fresh WT lines	<i>IND-2</i>	Isofemale Indiana, 9/97	Geldanamycin	W1	R. Krebs	9/97
	<i>IND-6</i>	Isofemale Indiana, 9/97	Geldanamycin	A1, W1, W2	R. Krebs	9/97
	<i>KYA-1D</i>	Kenya, 1996	<i>e1D, 9J1</i>	E5, W3	M. Ashburner	6/96
	<i>KYA-9C</i>	Kenya, 1996	<i>e1D, 9J1</i>	B1, L1, W1, W2	M. Ashburner	6/96
	<i>P-4</i>	Indiana, 8/96	<i>e1D, 9J1</i>	B1, B2, W3	S. Rutherford	8/96
	<i>P3-C</i>	Indiana, 8/96	<i>e6A, 9J1</i>	A1, B1	S. Rutherford	8/96
<i>Hsp83</i> alleles	<i>13F3</i>	R48C		L1	E. Hafen ¹⁵	12/94
	<i>19F2</i>	R48C		L1	E. Hafen ^{15,16}	12/94
	<i>9J1</i>	E377K		B2, L1, W1	E. Hafen ^{15,16}	12/94
	<i>e1D</i>	S38L		B2, E3, E4, E5, L1	G. Rubin ¹⁴	6/94
	<i>e3A</i>	S574C			G. Rubin ¹⁴	6/94
	<i>e4A</i>	S655F			G. Rubin ¹⁴	6/94
	<i>e6A</i>	S592F		A2	G. Rubin ¹⁴	6/94
	<i>e6D</i>	E317K			G. Rubin ¹⁴	6/94
	<i>P582</i>	P-element in 5' UTS		B2, E1, E2, L1, T1, W1	P. Deak	12/94
	<i>582[e11]</i>	Lethal P582 excision		W5	L. Yue	4/95

Strains producing deformed flies when crossed with Hsp90 mutants or raised on geldanamycin. Fly strains were divided into four categories: marked laboratory strains; wild-type laboratory stocks that had been maintained in the laboratory for many years; recently established (<1 year) wild-type lines; and lines carrying mutant *Hsp83* alleles. Mutant *Hsp83* alleles responsible for specific F₁ traits are indicated, with traits coded as in Table 1. Traits listed with the *Hsp83* alleles were observed in those stocks. UTS, untranslated sequence; WT, wild-type.

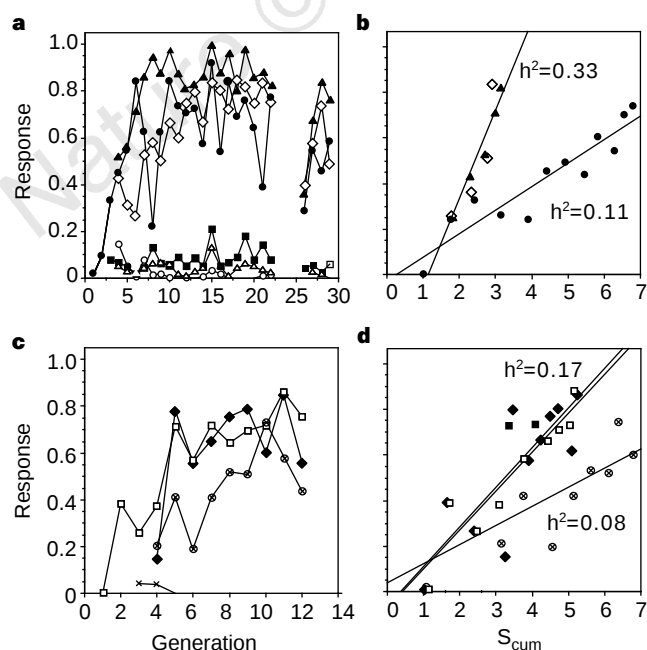


Figure 2 Selection experiments. **a, b**, Selection for the deformed-eye trait; **c, d**, Selection for the wing-vein trait. Selection lines: open circles, *LE1*; filled squares, *LE2*; open triangles, *LE3*; filled circles, *HE1*; open diamonds, *HE2*; filled triangles, *HE3*; crosses, *LV1*; open squares, *HV1*; circles with crosses, *HV2*; filled diamonds, *HV3*. *S_{cum}*, cumulative selection differential.

generation (Table 1). This pattern of heritability was unlikely to have been generated by *de novo* mutation and showed that the traits had a genetic basis, correlated to specific genetic backgrounds.

Production of specific traits

To determine whether the abnormal traits could be due to single genetic determinants interacting with the *Hsp83* mutations, we chose a wing and an eye trait for more extensive analysis. Neither trait was observed in any of the original parental stocks, but both appeared among the hybrid F₁ progeny of different crosses between the heterozygous *Hsp83* mutants and other stocks. One of these genetically hybrid F₁ progeny was a male with a deformed eye (Fig. 1m), arising from a cross between an *Hsp83* mutant (*19F2*) and a laboratory strain (*w¹¹¹⁸*) derivative. The other F₁ male had thickened wing veins (Fig. 1n) and was a hybrid resulting from a cross between an *Hsp83* line (*e6D*) and a wild-type strain (*Ives*). In both cases, when these males were crossed to a few (~5) unaffected F₁ females (*Hsp83*+) from the same parental cross, F₂ flies expressing a phenotype similar to that of the F₁ male were produced. Selection was initiated by crossing affected F₂ progeny to generate high-expression lines (high vein, HV, and high eye, HE). Unaffected progeny were crossed together to generate low-expression lines (LV and LE). We selected either for, or against, the trait in all subsequent generations. By the fourth generation, sufficient numbers of affected flies were produced to split each of the lines into three replicates.

The initial response to this selection regime indicated that the traits were polygenic (Fig. 2a, c). In successive generations, the proportion of affected progeny in the high-expression lines

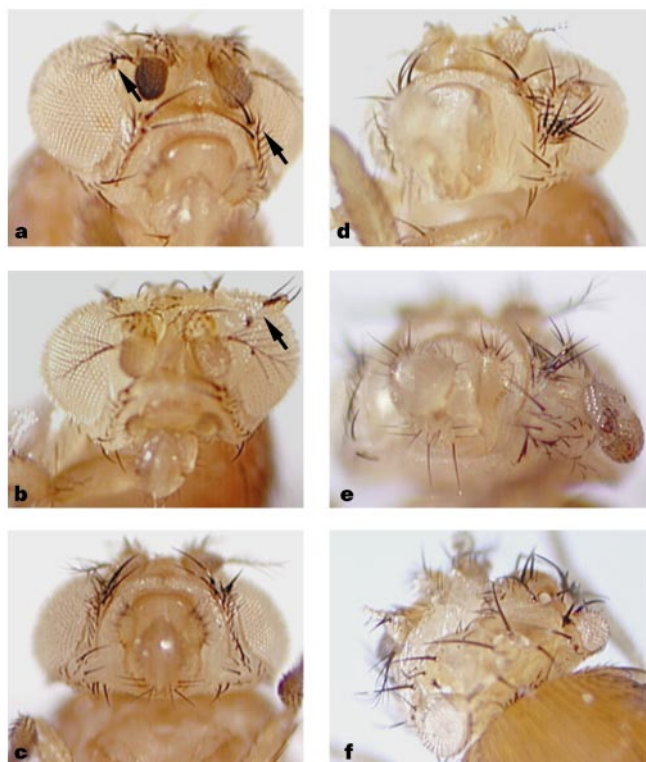


Figure 3 Variation within and between high-expression lines with the deformed-eye trait. **a–e**, Range of phenotypes in generation 14 of line *HE2*. **a**, Very mild duplications of wing margin bristles (arrows); **b**, small protuberances (arrow); **c–e**, severe bristle duplications and deformities. **f**, A severely affected animal from line *HE3*; in comparison with line *HE2*, many *HE3* mutants had small eyes and bare cuticle rather than overproduction of bristles. *HE3* was the only line in which flies with three antennae were produced.

increased and their phenotypes diverged. Furthermore, both traits exhibited a range of phenotypes within each line (Figs 3, 4). When the wing-vein trait was scored for increasing gradations of severity (Fig. 4), the distribution of affected flies varied markedly between the different high-expression lines (*HV1–3*) and between these lines and the base (parental) populations, which did not express the trait. All of the affected flies in line *HV2* expressed mild to moderate defects, whereas most of the affected flies from lines *HV1* and *HV3* expressed severe defects. This partitioning of the wing-vein phenotype between lines confirmed that more than one genetic determinant affected the trait.

The strong response to selection indicated that even though the founding populations were small, they contained a large amount of previously cryptic genetic variation that was capable of affecting these traits. We quantified this variation for each replicate line by regression of the initial selection response onto the cumulative selection differential¹⁸; the slope of each regression line is equal to its realized heritability (h^2 ; Fig. 2b, d). All h^2 values were non-zero, showing that selection had acted on pre-existing genetic factors influencing these traits.

Effect of temperature

Many of the traits seen in *Hsp83* heterozygotes were enhanced at either high (30°C) or low (18°C) temperatures (Table 1). For example, the wing-vein trait was produced in the high-expression lines that had been raised at 18°C but was not observed in replicate cultures grown at 25°C (results not shown). Conversely, the deformed-eye trait was enhanced at 30°C and was not observed at 18°C. To investigate this phenomenon further, we grew the six eye-

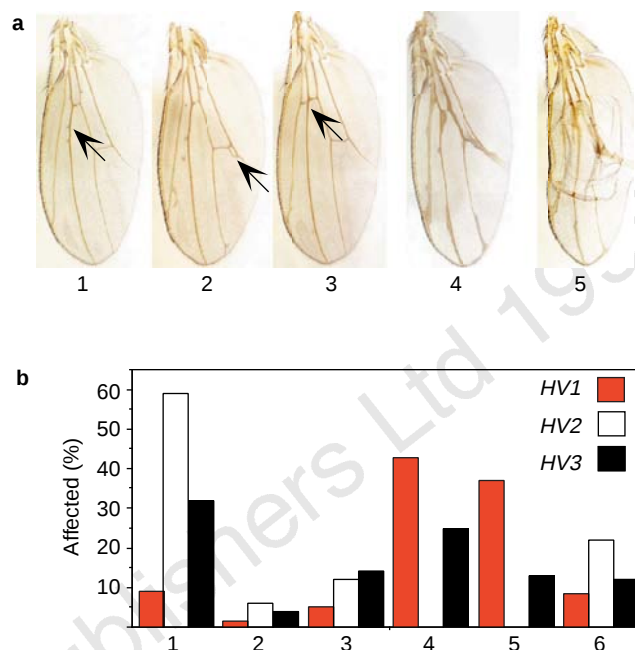


Figure 4 Partitioning of the wing-vein trait between lines. **a**, Wing defects of increasing severity from left to right: bumps on veins, looped veins, double veins, thick veins, and wing blisters. **b**, Per cent of affected individuals within each high-expression line (*HV1–3*) that exhibit traits of these severities (1–5), or other abnormal phenotypes (6).

trait selection lines at different temperatures within the normal growth range of *Drosophila* (Fig. 5). At 18°C the trait was not expressed. Between 23°C and 26°C, the penetrance and the severity of the trait in the three high-expression lines (*HE1–3*) increased sharply, whereas only a few flies from any of the low-expression lines (*LE1–3*) were affected. However, at 30°C, 20–30% of the flies in the low-expression lines presented the trait, indicating that some genetic determinants for the trait were still present in a cryptic state at the lower temperatures in these lines. The sigmoidal temperature–response curve (reaction norm; Fig. 5) indicates that

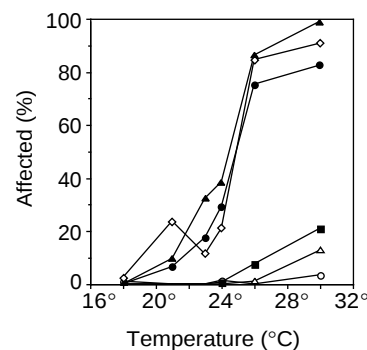


Figure 5 Temperature–response curves (norm of reaction) for the deformed-eye trait. Open circles, *LE1*; filled squares, *LE2*; open triangles, *LE3*; filled circles, *HE1*; filled triangles, *HE2*; open diamonds, *HE3*. Replicate cultures from generations 14–16 were grown at the indicated temperatures.

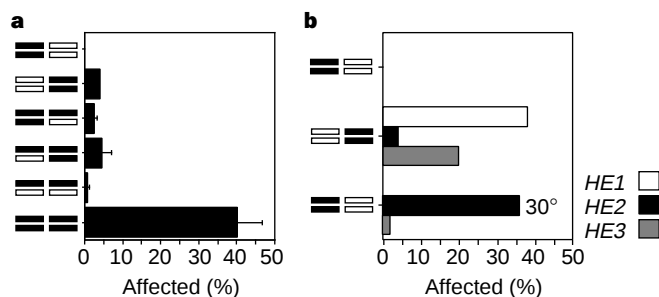


Figure 6 Genetic interactions in the *HE* lines. **a**, The percentage of flies affected ($\pm$ s.e.m.) by each of the six possible combinations of the two major autosomes (second and third) from line *HE2* (filled rectangles at left of graph) and the unselected control chromosomes (open rectangles). Second chromosomes are shown to the left and third chromosomes to the right. Similar results were obtained with *HE1* and *HE3* chromosomes. **b**, Second and third chromosomes from different *HE* lines differed in their ability to produce the trait at 18 °C or at 30 °C.

the eye phenotype is a threshold character¹⁸; the abrupt appearance of the trait depends on surpassing a threshold set by the number of underlying genetic determinants and their interaction with the environment.

Genetic interactions

To begin to characterize the genetic determinants underlying a particular morphological variant, we isolated chromosomes from each of the high-expression lines (*HE1-3*) in the context of a control genetic background. The effect of each chromosome in producing the deformed-eye phenotype was then tested in isolation and in combination with the other selected chromosomes. Figure 6 shows the complex relationship between the expression of the trait and the genotypes producing it. When either the second or the third chromosome from any of the high lines was heterozygous with the control chromosome in the new genetic background, the penetrance of the eye trait was markedly reduced (Fig. 6a). Homozygosity of any isolated second chromosome had no effect by itself at 25 °C, but appeared to affect the expression of genetic determinants on the third chromosome (Fig. 6). At 30 °C the isolated second chromosomes from line *HE2*, but not those from *HE1* or *HE3*, produced many affected flies (Fig. 6b). Thus, the second chromosomes in the three high-expression lines contain at least two different variants affecting the trait. When these determinants were separated from determinants present on the third chromosome (Fig. 6a), the expression of the trait disappeared.

Role of Hsp90

The deformed-eye lines were founded with *Hsp83* heterozygotes but selected only for the deviant trait, not the *Hsp83*-mutant chromosome (19F2). Indeed, two observations indicated that the trait might have become independent of the *Hsp83* mutation. First, even two extra copies of the wild-type *Hsp83* gene¹⁴ did not affect the expression of the trait when crossed into the high-expression lines (results not shown). Second, by the sixth and seventh generations, the trait was expressed in >80% of the flies in the *HE1* and *HE3* lines (Fig. 3a). If the trait continued to depend on the *Hsp83* mutation, as a result of mendelian segregation, at most two-thirds of the viable offspring of each generation would be *Hsp83* heterozygotes and potentially have deformed eyes. To determine directly whether the affected flies contained the 19F2 mutation (Arg 48 → Cys), we used polymerase chain reaction (PCR) genotyping (Fig. 7). In fact, none of the flies tested retained the mutation (generations 16–20; 0 out of 10 for each of lines *LE1*–3, *HE1* and *HE2*, and 0 out of 40 for line *HE3*). Therefore, selection had enriched the genetic determinants to a threshold at which

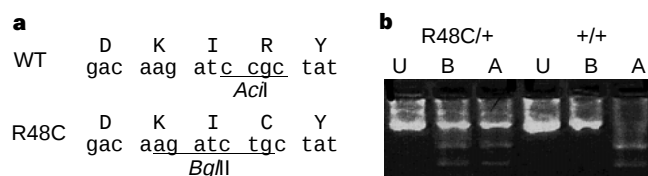


Figure 7 Genotyping of the 19F2 (R48C) *Hsp83* mutation. **a**, The base change causing the R48C mutation destroys an *AccI* restriction site present in the wild-type sequence (WT) and creates a *Bgl*II site. **b**, Uncut controls (U) are 575 base pairs long. Characteristic *Bgl*II (B) or *AccI* (A) digests, producing 335- and 240-base-pair products, are shown for R48C heterozygotes (R48C/+) and wild-type flies (+/+).

the expression of the trait became independent of the Hsp90 mutation.

When the high-expression lines were outcrossed with normal laboratory strains, the deformed-eye trait was expressed only at very low levels. We took advantage of this fact to re-examine the role of Hsp90 in buffering these now cryptic determinants. The high-expression lines were backcrossed to the parental *Hsp83*-mutant stock (19F2/TM6B), which contained one mutant and one wild-type copy of *Hsp83*. At all temperatures the trait exhibited low penetrance in the progeny receiving the control chromosome (TM6B; Fig. 8, filled diamonds). The progeny receiving the *Hsp83*-mutant (19F2) expressed the trait at a significantly higher level (Fig. 8, open squares). At lower temperatures, or when supplied with normally functioning Hsp90, determinants for the trait were kept in a cryptic state, but there was a dramatic increase in expression of the trait at temperatures above 25 °C. Similar temperature-response curves were produced in crosses between flies containing this mutant *Hsp83* allele and the other high-expression lines (*HE1* and *HE3*). Moreover, another independently isolated *Hsp83* allele, 13F3, also enhanced the expression of the trait. Thus, wild-type Hsp90 buffered the penetrance and temperature sensitivity of the trait, coupling the appearance of this morphogenetic variant to the environment.

Natural variation in wild flies

Laboratory stocks might differ from wild flies in many respects: genetic drift, inbreeding, relaxed selection and marker mutations might decrease the developmental stability of typical laboratory strains. To determine whether wild flies also carry the same types of morphogenic variation, we obtained previously established wild-type lines from several geographic locations and collected wild flies from the field. When these flies were crossed to *Hsp83*-mutant heterozygotes, morphological abnormalities appeared in F₁ hybrids, both in the recently established wild-type lines and in wild-type laboratory stocks originating from globally diverse locations (Table 1). Nearly 25% of the crosses with wild-type laboratory stocks (37 of 163 crosses, using at least 20 different wild-type laboratory stocks) produced abnormal flies, numbering 120 out of 16,080 scored (0.8%). Two of eleven recently established lines from a wild population represented by independent isofemale lines (*IND-1–12*) produced 8% deformed flies when treated with geldanamycin (45 of 559 flies studied; Table 1). In fact, several flies from line *IND-6* had notched wings (Fig. 1p), while many flies with small wings on one or both sides appeared in line *IND-2*, although neither defect was found in the untreated stocks. Therefore, silent polymorphisms buffered by *Hsp90* must also exist in the wild.

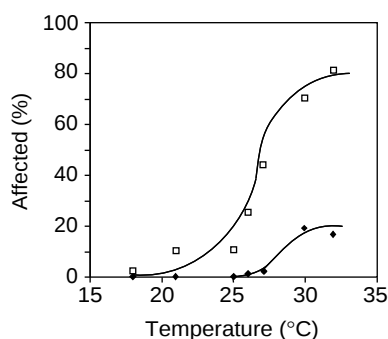


Figure 8 Wild-type Hsp90 buffered the deformed-eye trait. An *Hsp83* mutant stock (19F2/TM6B) was crossed to the HE2 high-expression line and replicate cultures were grown at different temperatures. Results from progeny containing the 19F2 chromosome (open squares) and from controls containing wild-type *Hsp83* on chromosome TM6B (filled diamonds) are shown.

A mechanism for 'evolvability'?

We have provided what is, to our knowledge, the first evidence for an explicit molecular mechanism that assists the process of evolutionary change in response to the environment. We suggest that in nature, transient decreases in Hsp90 levels resulting from its titration by stress-damaged proteins could uncover morphological variants for selection to act upon. Consider a simple model for a threshold trait requiring at least six genetic determinants (with no dominance). In a population containing ten independent and additive determinants affecting the trait, each present at a frequency of 0.1, the probability of an individual having at least six of these determinants and thus the trait, is about 1 in 7,000. However, if compromising Hsp90 function were to lower the trait's threshold by just one or two determinants, the probability of the appearance of the trait increases to 1 in 600 or 1 in 78. Once the frequency of a trait is increased in this manner, given a moderate fitness advantage, selection could increase the frequency of genetic polymorphisms affecting the trait to a point at which it no longer depends on reduced Hsp90 function to be expressed in the population.

Evolutionary models must encompass a dichotomy of stasis and change. Evolution exploits genetic differences between individuals in order to remodel developmental programs, yet development is generally robust to individual genetic differences and environmental perturbations. Theoretical models describe how developmental homeostasis is developed and why it is maintained, as well as how it could be disrupted so that evolutionary change can occur^{19–23}. Experiments disrupting developmental homeostasis by specific mutations^{24–26}, by particular stresses during precise windows of development (phenocopy)^{27,28}, or in species hybrids²⁹, have shown that populations contain a surprising amount of unexpressed genetic variation that is capable of affecting certain typically invariant traits. Sometimes very specific conditions can uncover this previously silent variation^{30–32}. But both the wide variety and unusual character of the morphological variation uncovered when Hsp90 is impaired, and the prevalence of natural stresses that might disrupt it, are unprecedented.

Drosophila Hsp83 mutants have been isolated repeatedly in near-saturation genetic screens for genes that interact in signal transduction^{14–16,33}. Mutations in other components of the protein-folding and translation machinery have not been identified in such screens. Moreover, known mutations that generally affect protein biogenesis in *Drosophila*, such as mutations in ribosomal subunits, do not produce such abnormal morphologies³⁴. Hsp90 is special because its participation in the stress response is coupled with its critical integrative position in the genetic architecture of development. For example, it seems likely that many variant morphologies could be produced only by changing the output of

several developmental processes simultaneously. Genetically 'sensitized' pathways, engineered to destabilize certain phenotypes, reveal potent natural variation affecting these phenotypes²⁶. By altering the activities of multiple signal transducers and thereby simultaneously weakening several developmental pathways, Hsp90 can expose such variation, allowing selection to remodel many different processes at once.

In some cases, such as cell-cycle control^{10,14,35}, Hsp90 supports both the activators and the inhibitors of the same function and alteration of Hsp90 function could uncover variation that would allow selection to sculpt the output of these processes either upwards or downwards. Furthermore, this mechanism is flexible. When Hsp90 function is disturbed, developmental pathways are sensitized to a degree determined by their specific dependence on Hsp90 (which is dictated by the functional significance and inherent stabilities of the relevant targets) and by Hsp90 availability (which is dictated in nature by the severity of the stress). The use of Hsp90 as a capacitor for the conditional release of stores of hidden morphogenic variation may have been adaptive for particular lineages, perhaps allowing the rapid morphological radiations that are found in the fossil record. □

Methods

Drosophila strains and culture. Strains used in this study are listed in Table 2. The *Hsp83* alleles 19F2 and 13F3 are independent isolates of the same point mutation (E. Hafen, personal communication). *IND* isofemale lines were established from larvae found in separate pieces of fruit to decrease the probability that the lines were started from closely related flies. Flies were cultured on standard cornmeal:molasses:agar *Drosophila* medium, except in inhibitor experiments, in which 3 µg ml⁻¹ geldanamycin (National Cancer Institute) was made up in a 1% bakers yeast slurry before an equal volume of Formula 4-24 Instant *Drosophila* Medium (North Carolina Biological Supply) was added.

Quantitative analyses. Responses to selection (Fig. 2) were measured as the fraction of affected progeny in each generation; each fly was given a trait value of 0 (not affected) or 1 (affected). We used the difference in the selection response between successive generations to determine the selection differential. Realized heritability (h^2 ; Fig. 2b, d) was obtained from the slope of the least-squares regression of selection response plotted against the cumulative selection differential (S_{cum})¹⁸. Genotype–environment interactions ($G \times E$) were confirmed for the data shown in Fig. 5 by analysis of variance (SAS Inst.). There were highly significant interaction effects ($G \times E$) for the high- and low-expression lines ($P < 0.0001$); and for these lines compared with wild-type lines, which are unaffected at all temperatures ($P < 0.0004$). Genic interaction effects due to epistasis are suggested by the data in Fig. 6. The difference apparent from comparison of the sum of the genotypic values for the isolated second and third chromosomes compared with their combined genotypic value is the interaction deviance due to epistasis between these chromosomes¹⁸. Epistasis can arise from different underlying causes, including an expression threshold for additive determinants³⁶; however, when these data are transformed to a different scale suggested for use with threshold traits (deviation from mean liability in units of standard deviation)¹⁸ a non-additive interaction, indicating more general epistasis, is still apparent (results not shown).

PCR genotyping. Convergent primers framing a 575-base-pair region surrounding the mutation were used to amplify by PCR genomic DNA isolated from single flies by proteinase-K digestion. Digestion products were resolved by electrophoresis through 2% NuSieve agarose (FMC Bioproducts) and were visualized with ethidium bromide.

Received 15 July; accepted 1 October 1998.

- Morimoto, R. I., Kline, M. P., Bimston, D. N. & Cotto, J. J. The heat-shock response: regulation and function of heat-shock proteins and molecular chaperones. *Essays Biochem.* **32**, 17–29 (1997).
- Nathan, D. F., Vos, M. H. & Lindquist, S. *In vivo* functions of the *Saccharomyces cerevisiae* Hsp90 chaperone. *Proc. Natl Acad. Sci. USA* **94**, 12949–12956 (1997).
- Rutherford, S. L. & Zuker, C. S. Protein folding and the regulation of signaling pathways. *Cell* **79**, 1129–1132 (1994).
- Smith, D. F. Dynamics of heat shock protein 90–progesterone receptor binding and the disactivation loop model for steroid receptor complexes. *Mol. Endocrinol.* **7**, 1418–1429 (1993).
- Nathan, D. F. & Lindquist, S. Mutational analysis of Hsp90 function: interactions with a steroid receptor and a protein kinase. *Mol. Cell. Biol.* **15**, 3917–3925 (1995).

6. Jakob, U., Lilie, H., Meyer, I. & Buchner, J. Transient interaction of Hsp90 with early unfolding intermediates of citrate synthase. Implications for heat shock *in vivo*. *J. Biol. Chem.* **270**, 7288–7294 (1995).
7. Freeman, B. C. & Morimoto, R. I. The human cytosolic molecular chaperones hsp90, hsp70 (hsc70) and hsp110 have distinct roles in recognition of a non-native protein and protein refolding. *EMBO J.* **15**, 2969–2979 (1996).
8. Picard, D. *et al.* Reduced levels of hsp90 compromise steroid receptor action *in vivo*. *Nature* **348**, 166–168 (1990).
9. Holley, S. J. & Yamamoto, K. R. A role for Hsp90 in retinoid receptor signal transduction. *Mol. Biol. Cell.* **6**, 1833–1842 (1995).
10. Stepanova, L., Leng, X., Parker, S. B. & Harper, J. W. Mammalian p50Cdc37 is a protein kinase-targeting subunit of Hsp90 that binds and stabilizes Cdk4. *Genes Dev.* **10**, 1491–1502 (1996).
11. Xu, Y. & Lindquist, S. Heat-shock protein hsp90 governs the activity of pp60v-src kinase. *Proc. Natl Acad. Sci. USA* **90**, 7074–7078 (1993).
12. Ali, A., Bharadwaj, S., O'Carroll, R. & Ovsenek, N. HSP90 interacts with and regulates the activity of heat shock factor 1 in *Xenopus* oocytes. *Mol. Cell. Biol.* **12**, 4949–4960 (1998).
13. Zou, J., Guo, Y., Guettouche, T., Smith, D. F. & Voellmy, R. Repression of heat shock transcription factor HSF1 by HSP90 (HSP90 complex) that forms a stress-sensitive complex with HSF1. *Cell* **94**, 471–480 (1998).
14. Cutforth, T. & Rubin, G. M. Mutations in Hsp83 and cdc37 impair signaling by the sevenless receptor tyrosine kinase in *Drosophila*. *Cell* **77**, 1027–1036 (1994).
15. Dickson, B. J., van der Straten, A., Dominguez, M. & Hafen, E. Mutations modulating Raf signaling in *Drosophila* eye development. *Genetics* **142**, 163–171 (1996).
16. van der Straten, A., Rommel, C., Dickson, B. & Hafen, E. The heat shock protein 83 (Hsp83) is required for Raf-mediated signalling in *Drosophila*. *EMBO J.* **16**, 1961–1969 (1997).
17. Whitesell, L., Mimnaugh, E. G., De Costa, B., Myers, C. E. & Neckers, L. M. Inhibition of heat shock protein HSP90-pp60v-src heteroprotein complex formation by benzoquinone ansamycins: essential role for stress proteins in oncogenic transformation. *Proc. Natl Acad. Sci. USA* **91**, 8324–8328 (1994).
18. Falconer, D. S. & Mackay, T. F. C. *Introduction to Quantitative Genetics* (Longman, Harlow, 1996).
19. Waddington, C. H. Canalization of development and the inheritance of acquired characters. *Nature* **150**, 563–565 (1942).
20. Wagner, G. P. Coevolution of functionally constrained characters: prerequisites for adaptive versatility. *Biosystems* **17**, 51–55 (1984).
21. Kauffman, S. A. *The Origins of Order: Self Organization and Selection in Evolution* (Oxford Univ. Press, New York, 1993).
22. Raff, R. A. *The Shape of Life: Genes, Development, and the Evolution of Animal Form* (Univ. Chicago Press, Chicago, 1996).
23. Gerhart, J. & Kirschner, M. *Cells, Embryos, and Evolution: Toward a Cellular and Developmental Understanding of Phenotypic Variation and Evolutionary Adaptability* (Blackwell, Malden, 1997).
24. Waddington, C. H. Genetic assimilation of an acquired character. *Evolution* **7**, 118–126 (1953).
25. Rendel, J. M., Sheldon, B. L. & Finlay, D. E. Canalisation of development of scutellar bristles in *Drosophila* by control of the scute locus. *Genetics* **52**, 1137–1151 (1965).
26. Polaczyk, P. J., Gasperini, R. & Gibson, G. Naturally occurring genetic variation affects *Drosophila* photoreceptor determination. *Dev. Genes Evol.* **207**, 462–470 (1998).
27. Waddington, C. H. Genetic assimilation of the *bithorax* phenotype. *Evolution* **10**, 1–13 (1956).
28. Gibson, G. & Hogness, D. S. Effect of polymorphism in the *Drosophila* regulatory gene Ultrabithorax on homeotic stability. *Science* **271**, 200–203 (1996).
29. Wade, M. J., Johnson, N. A., Jones, R., Siguel, V. & McNaughton, M. Genetic variation segregating in natural populations of *Tribolium castaneum* affecting traits observed in hybrids with *T. freemani*. *Genetics* **147**, 1235–1247 (1997).
30. De Jong, G. & Scharloo, W. Environmental determination of selective significance or neutrality of amylase variants in *Drosophila melanogaster*. *Genetics* **84**, 77–94 (1976).
31. Dykhuizen, D. & Hartl, D. L. Selective neutrality of 6PGD allozymes in *E. coli* and the effects of genetic background. *Genetics* **96**, 801–817 (1980).
32. Kimura, M. *The Neutral Theory of Molecular Evolution* (Cambridge Univ. Press, Cambridge, New York, 1983).
33. Simon, M. A., Bowtell, D. D., Dodson, G. S., Lavery, T. R. & Rubin, G. M. Ras1 and a putative guanine nucleotide exchange factor perform crucial steps in signaling by the sevenless protein tyrosine kinase. *Cell* **67**, 701–716 (1991).
34. Lindsley, D. L. & Zimm, G. G. *The Genome of Drosophila melanogaster* (Academic, San Diego, 1992).
35. Aligue, R., Akhavan-Niak, H. & Russell, P. A role for Hsp90 in cell cycle control: Wee1 tyrosine kinase activity requires interaction with Hsp90. *EMBO J.* **13**, 6099–6106 (1994).
36. Futuyma, D. J. *Evolutionary Biology* (Sinauer, Sunderland, 1998).
37. Xu, Y., Singer, M. & Lindquist, S. Maturation of c-src as a kinase and as a substrate is dependent on Hsp90. *Proc. Natl Acad. Sci. USA* (in the press).

Acknowledgements. We thank M. Wade for discussions, J. Segre for assistance in devising the PCR genotyping strategy, laboratories that supplied fly strains (see Table 2), and T. Mackay for insight into the quantitative analysis of threshold traits.

Correspondence and requests for materials should be addressed to S.L.

Protostellar jets irradiated by massive stars

Bo Reipurth*, John Bally*, Robert A. Fesen† & David Devine‡

* Center for Astrophysics and Space Astronomy, Campus Box 389,
University of Colorado, Boulder, Colorado 80309, USA

† Dartmouth College, Department of Physics and Astronomy, Hanover,
New Hampshire 03755, USA

‡ Laboratory for Astronomy and Solar Physics, Code 681,
NASA Goddard Space Flight Center, Greenbelt, Maryland 20771, USA

The formation of solar-type stars is a gradual process during which they accrete mass from the dense disks and cloud cores that surround them. This accretion requires the release of angular momentum, and an important mechanism for achieving this seems to be the production of jets along the polar axes of the young stars^{1,2}. But the presence of massive, luminous stars within the same star-forming region can affect the forming stars by stripping away their circumstellar envelopes with ultraviolet radiation, thereby removing the reservoir of gas from which the stars are built up and exposing the disks to photoerosion³. Here we present observations of four highly collimated jets from young stars that appear to have been stripped of their circumstellar molecular cloud cores in this way. The production of jets seems to have been largely unaffected. If these jets are also photoionized, their mass loss rates can be determined from observations with much greater accuracy than for normal shock-excited jets.

The young multiple star σ Orionis, located just below Orion's Belt, has five massive components, the most luminous being an O9.5 star⁴ at least 15 times more massive than our Sun. These stars excite a large and tenuous H II region bordered on the eastern side by a wall of neutral material from which the Horsehead nebula protrudes. σ Orionis is part of the Orion 1b association, with an age variously determined to be 1.7, 5.1 and 7 Myr (refs 5–7) and a distance between 360 and 470 pc (refs 5, 8). Almost 100 H α emission stars⁹ and/or X-ray sources¹⁰ have been found around σ Orionis which represent a young low-mass stellar population that may have formed about 2 million years ago¹¹.

We have obtained narrow-band images of the region between σ Orionis and the Horsehead nebula with the new 8192 $\times$ 8192 pixel MOSAIC imager on the Mayall 4-m telescope at Kitt Peak National Observatory near Tucson, Arizona (see Fig. 1 legend for details). We discovered four collimated Herbig–Haro (HH) jets which we here name HH444–HH447. HH jets are highly collimated supersonic outflows from newborn stars which are powered by episodic mass-accretion events from circumstellar disks^{12,13}. Two of these jets, HH444 and HH445, have extended lobes containing more distant shocks, which in the case of the former jet have clear bow shock morphologies (Fig. 1). The detailed structure of the four jets are shown in Fig. 2, and individual characteristics are discussed in Fig. 2 legend. The four flows share several traits: first, they are all highly asymmetric with at most poorly developed counterjets. Second, the driving sources of all four jets are visible, relatively bright stars; whereas most jet sources in other star-forming regions are shrouded from view at optical wavelengths by opaque clouds of gas and dust, all four jets reported here are driven by visible stars with low reddening. Third, all four jets are within a few parsecs of σ Orionis and are thus bathed in the ultraviolet radiation emitted by these massive stars, and are among the first externally irradiated jets to be recognized¹⁴.

We also obtained spectra of the two brightest HH jets and of all four sources (some are shown in Fig. 3) with the Hiltner 2.4-m telescope at the Michigan–Dartmouth–MIT Observatory located

on Kitt Peak. The spectra of these two jets differ from those of purely shock-excited HH jets. Instead, they resemble spectra normally emitted by photoionized gas, indicating that much of the radiating gas in these jets is made visible not by internal shocks but by the ultraviolet radiation field of the nearby massive star σ Orionis. This suggests the possibility of an unambiguous determination of the jet densities and mass loss rates. In all other stellar jets, these parameters have to be extracted from highly uncertain models of the shocks which make them visible.

All of the source stars have bright H α emission lines, and their spectra suggest that they are T Tauri stars. None of the stellar spectra show continua rising steeply towards the red, indicating that the stars are not heavily reddened. None of them were detected by the IRAS satellite, possibly because they have very little circumstellar material. The source of the most prominent jet, HH444, is a variable star with a visual magnitude around 14 to 16 known as V510 Orionis^{9,15–17}. Earlier spectra of this star provided preliminary evidence for outflow activity¹⁵ and for active accretion¹⁶. The driving sources of the other jets have visual magnitudes of about 15 or fainter. Background stars as well as some galaxies are seen around these stars, emphasizing the absence of significant molecular material. A 2.6-mm-wavelength observation of the $^{12}\text{CO } J = 1 \rightarrow 0$

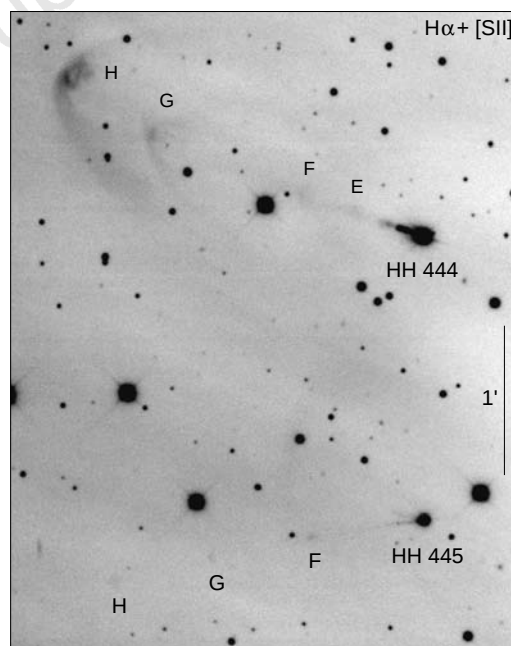


Figure 1 The two jets HH444 and HH445. This image was formed by combining two exposures through 80-Å-wide H α and [S II] filters for a total exposure time of 1,500 s in each filter. HH444 is a bright jet emanating from a visible star with a prominent 16"-long beam that points towards a large diffuse bow shock (H) with its apex 154" from the source. Two fainter bow shocks (G, F) are seen ~114" and 52" from the source, indicating multiple eruptions of the source. The three bow shocks have slightly decreasing position angles with distance from the source, suggesting that the outflow axis is gradually changing position with time. The total projected extent of the HH444 complex is ~0.35 pc, a size common for HH flows. If we assume a tangential flow velocity of about 150 km s⁻¹, typical of HH flows, the kinematic age of the flow complex is ~2,200 yr. About 2" south of HH444, we find a second faint but well collimated jet, which we call HH445, emanating from a fairly bright visible star. Further along the well defined flow axis there are three faint knots F, G and H about 45", 86" and 125" from the source, respectively. Again, a slight curvature of the flow axis is apparent. At an assumed distance of 470 pc, the total projected extent of the HH445 flow is 0.28 pc. Assuming a tangential velocity of 150 km s⁻¹, knots F, G and H are about 600, 1,200 and 1,800 yr old, suggesting periodic activity of the source. The large number of background stars testifies to the very small amount of interstellar material in the region.

transition with the Swedish–European Southern Observatory 15-m dish at La Silla in Chile towards V510 Orionis reveals two very faint lines, of which at least one must be a background signal. Thus, these four stars show none of the usual signatures of abundant circumstellar material. However, as the stars have jets they probably do have disks, but so small (perhaps less than 1" in diameter) that sub-arcsecond observing techniques are required to detect them. The apparent absence of significant circumstellar material is consistent with a 2-million-year age estimated for the σ Orionis sub-group of the Orion 1b association¹¹.

If the stars are indeed this old, then jet production in young stars must continue for at least a few million years. Vestiges of the jet phenomenon from what appears to be more-evolved T Tauri stars in quiescent regions like Taurus have been observed as microjets^{18–20}. The nature of the reservoir that fuels such microjets as well as the present irradiated jets has to be explained, because at least the latter do not come from stars associated with extended cloud cores. One possibility is that the driving sources are unresolved binary stars with high-eccentricity orbits²¹. Strong tidal forces during periastron passage could trigger accretion events in remnant circumstellar disks²² to fuel and trigger episodic jet formation. The twin jets in

HH445 (Fig. 2) might provide indirect support for the binary-star model. Another possibility is that these stars have formed their first protoplanets which are spiralling into their central stars due to interactions with left-over disk material. Although the energy release would be sufficient to drive a jet, it is unclear how a jet could be launched or collimated by a protoplanet accretion event. Finally, if jets are magnetically launched and collimated, photo-evaporation of the disk could actually help load material onto the magnetic field lines to revive an otherwise fading outflow.

Alternatively, it is possible that the low-mass stars near σ Orionis were not all formed at the same time but have an age gradient, as seen in various other star-forming regions²³. The four jet sources may be the very last of the group to form, and could therefore be much younger than 2 million years. In this scenario, the massive σ Orionis stars could have had time to clear away the molecular material usually abundant in star-forming regions, thus leaving the jet sources in their present naked state, mimicking the appearance of much older stars.

In the absence of external effects, it is expected that very young jets should be symmetric. The asymmetry seen in many of the best known HH jets like HH34, HH47 and HH111^{24–26} is ascribed to heavy obscuration of the receding lobe, and indeed infrared studies often reveal intrinsic symmetry in these jets^{27,28}. However, this explanation does not work for the jets discussed here, because they are not associated with extended high-extinction regions. More evolved microjets are often quite asymmetric²⁹, and also the four irradiated jets show dramatic differences between their lobes, perhaps for the same (unknown) reasons. But it is perhaps noteworthy that, in projection, all four jets point away from the OB stars, which for most geometries implies that the poorly developed jet lobes are the ones that are most exposed to strong ultraviolet

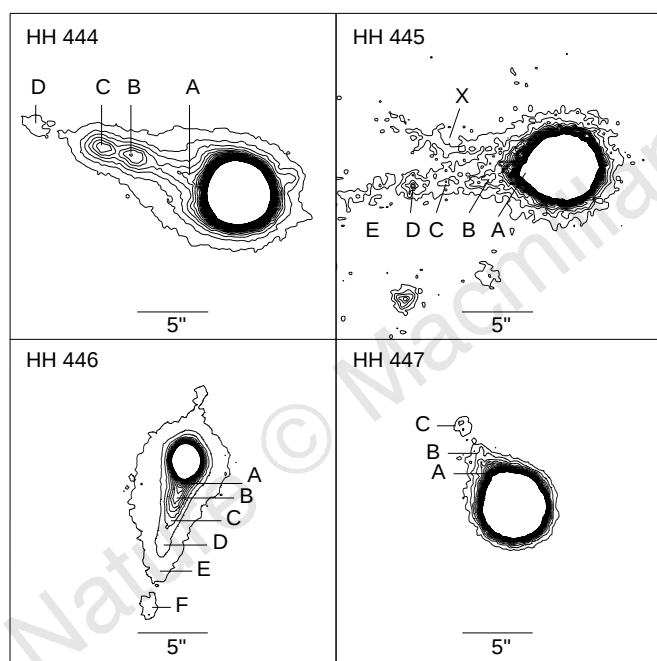


Figure 2 The inner jets and their stellar sources, as seen in logarithmic contour plots. Each figure is 23×23 arcsec. **a**, HH444 (source located at right ascension (RA) 5 h 37 min 9.0 s, declination (dec.) $-2^{\circ} 32' 56''$ (in 1950 coordinates), known as V510 Ori). The contour diagram shows that the well collimated jet consists of four diffuse knots (A–D), with B and C much brighter than the others. The $H\alpha$ and [S II] images show a strong gradient in the $H\alpha$ /[S II] emission ratio along the jet, with knot A being very strong in [S II] and knot D very strong in $H\alpha$. A very faint and rudimentary counterjet can also be discerned. **b**, HH445 (source located at RA 5 h 37 min 9 s, dec. $-2^{\circ} 34' 50''$ (1950 coord.), known as A0976-357). The HH445 jet contains several faint knots (A–E) and is 18" long. Closer examination reveals a second jet, which we call HH445X, slightly fainter and more diffuse, emanating from the same source. This highly unusual twin jet system may indicate that the source is a close binary star. Of the two, only the HH445A jet has a counterjet, but as in the case of HH444, it is extremely faint. **c**, HH446 (source located at RA 5 h 36 min 54.9 s, dec. $-2^{\circ} 40' 20''$ (1950 coord.)). The HH446 jet lies $\sim 8'$ southwest of HH444, emerges from a visible star and stretches at least 11" (knots A–F) towards the south. Again, only a very faint counterjet that is a few arc seconds long is visible. **d**, HH447 (source located RA 5 h 37 min 38.1 s, dec. $-2^{\circ} 35' 5''$ (1950 coord.), known as Haro 5-39). This fourth jet is almost $8'$ to the east of HH444, and is only $\sim 7''$ long. This jet is also associated with a visible star and there is not even a hint of a counterjet.

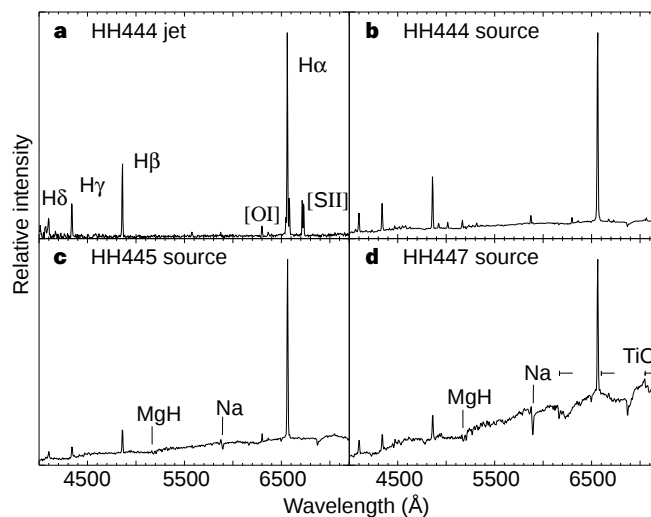


Figure 3 Background-subtracted spectra of a jet and three sources. **a**, The HH444 jet (knots B and C) is dominated by Balmer emission lines that are unusually strong relative to the [O I] and [S II] lines. A velocity difference of 220 km s^{-1} was measured between the main (blue-shifted) lobe and the faint counterlobe. A spectrum of the bow shock (H, not shown) has strong [O III] emission. A spectrum of the HH446 jet shows similar lines, but even stronger Balmer lines and a much higher electron density; the two other jets were too faint to observe spectroscopically. **b**, The source of HH444 (V510 Ori) shows a heavily veiled continuum with a rich emission-line spectrum. **c**, The source of HH445 (A0976-357) shows a K-type stellar spectrum with MgH absorption at $5,200 \text{ Å}$, prominent Balmer emission lines and no evidence for significant reddening. **d**, The source of HH447 (Haro 5-39) shows an M-type stellar spectrum with prominent TiO bands, strong sodium absorption at $5,893 \text{ Å}$, only Balmer lines in emission, and again no evidence for any significant reddening beyond the intrinsically red stellar continuum. All these sources are likely to be T Tauri stars. All images and spectra presented here were obtained in October 1997.

radiation. On the one hand, as noted above, the increased photoevaporation of the disk surface might load material onto magnetic field lines, which should produce a heavier jet. On the other hand, the role that strong radiation fields have on the region of jet collimation is unstudied, and it is conceivable that formation of a normal jet might be impeded on the illuminated side. Although the well developed jet is also exposed, the point of launch and collimation may lie in the shadow of a small circumstellar disk that remains undetected in our data and is thus only illuminated by the diffuse ultraviolet field. The best-developed jet, HH444, shows a normal shock-excited [S II] bright region at the base of the jet, which rapidly changes to an H α dominated seemingly photoionized spectrum further away from the driving source.

If these jets are photoionized, this opens the possibility of determining fundamental jet parameters like the mass-loss rate much more accurately than is possible for common shock-excited jets. A jet with a steady, constant mass loss rate and a sound speed c_s will widen at an angle given by $2c_s/v_j$ where v_j is the jet velocity. Therefore, the jet density declines as r^{-2} , where r is the distance from the jet source. As the jet emerges from the shadow of the small disk associated with its source star, ultraviolet radiation will rapidly ionize the skin of the jet until the total recombination rate in the photoionized zone just balances the incident ionizing flux. In this region the jet contains both a neutral core and a photoionized exterior. However, the divergence of the jet ensures that it eventually becomes fully ionized as it moves beyond a distance given by $r_{\max} = 4.6 \times 10^{16} \sin^{-1/3}(\theta) \dot{M}_{-8}^{2/3} V_{200}^{1/3} c_{10}^{-1} Q_{48}^{-1/3} d_3^{-1}$ cm, where $\dot{M}_{-8}$ is the jet mass-loss rate in units of 10^{-8} solar masses per year, V_{200} is the jet velocity in units of 200 km s^{-1} , c_{10} is the sound speed in units of 10 km s^{-1} , Q_{48} is the number of ionizing photons emitted by σ Orionis in units of 10^{48} photons s^{-1} , d_3 is the distance between σ Orionis and the jet in units of 3 pc, and θ is the angle between the jet axis and the arrival direction of the ionizing photons. (We assume that the jet density at any particular distance from its source star is uniform across the jet, and we ignore the diffuse radiation field). An irradiated jet with the above parameters set to unity is expected to become fully ionized when the density drops below about 600 cm^{-3} , which is the jet density at $r_{\max} = 4.6 \times 10^{16} \text{ cm}$ (equivalent to $7''$ to $9''$ at 360 to 470 pc). This is in good agreement with observations of HH444, where [O I] emission tracing neutral material extends for about $8''$ to $10''$ from V510 Orionis (knots B and C). The [S II] line ratio in knots B and C indicates an electron density n_e of about $330 \pm 30 \text{ cm}^{-3}$. If the jet becomes fully ionized beyond this point, then this density should agree with that which can be derived from the H α surface brightness (proportional to the emission measure $n_e^2 l$ where l is the thickness of the jet); this implies that the jet is about 450 AU in diameter (about $1''$) and transports mass at a rate of about $5 \times 10^{-8}/V_{200}$ solar masses per year through knots B and C. More detailed future observations of irradiated jets could be used to measure directly the density and mass in a jet as a function of distance from its source, leading to better determinations of mass loss rates from young stars. $\square$

Received 30 April; accepted 4 September 1998.

- Shu, F. H., Najita, J., Wilkin, F., Ruden, S. P. & Lizano, S. Magneto-centrifugally driven flows from young stars and disks. I. A generalized model. *Astrophys. J.* **429**, 781–796 (1994).
- Ouyed, R., Pudritz, R. E. & Stone, J. Episodic jets from black holes and protostars. *Nature* **385**, 409–414 (1997).
- Johnstone, D., Hollenbach, D. & Bally, J. Photoevaporation of disks and clumps by nearby massive stars: application to disk destruction in the Orion Nebula. *Astrophys. J.* **499**, 758–776 (1998).
- Hoffleit, D. *The Bright Star Catalogue* (Yale Univ., New Haven, CT, 1982).
- Brown, A. G. A. *Stellar Content and Evolution of OB Associations*. Thesis, Univ. Leiden (1996).
- Warren, W. H. & Hesser, J. E. A photometric study of the Orion OB1 association. III-Subgroup analyses. *Astrophys. J. Supp.* **36**, 497–572 (1978).
- Blaauw, A. in *The Physics of Star Formation and Early Stellar Evolution* (eds Lada, C. J. & Kylafis, N. D.) 125–154 (NATO ASI Ser. C, Vol. 342, Kluwer, Dordrecht, 1991).
- de Zeeuw, P. T., Hoogerwerf, R., de Bruijne, J. H. J., Brown, A. G. A. & Blaauw, A. A Hipparcos census of the nearby OB associations. *Astron. J.* (in the press).
- Wiramihardja, S. D., Kogure, T., Yoshida, S., Ogura, K. & Nakano, M. Survey observations of emission-line stars in the Orion region. I. The Kiso area A-0904. *Publ. Astron. Soc. Jpn* **41**, 155–174 (1989).
- Wolk, S. J. *Watching the Stars Go 'Round and 'Round*. Thesis, State Univ. New York at Stony Brook (1996).

- Walter, F. M., Wolk, S. J. & Sherry, W. The Sigma Orionis Cluster. In *Cool Stars, Stellar Systems, and the Sun 10* (eds Donahue, R. & Bookbinder, J.) (ASP Conf. Ser., in the press).
- Reipurth, B. & Heathcote, S. in *Herbig-Haro Flows and the Birth of Low Mass Stars* (eds Reipurth, B. & Bertout, C.) 3–18 (IAU Symp. 182, Kluwer, Dordrecht, 1997).
- Hartmann, L. in *Herbig-Haro Flows and the Birth of Low Mass Stars* (eds Reipurth, B. & Bertout, C.) 391–404 (IAU Symp. 182, Kluwer, Dordrecht, 1997).
- Cernicharo, J., LeFloch, B., García López, R. & Esteban, C. in *Low Mass Star Formation—from Infall to Outflow* (eds Malbet, F. & Castets, A.) 8–10 (Observ. de Grenoble, 1997).
- Sanduleak, N. Detection of T Tauri-like spectra in six stars in Orion. *Publ. Astron. Soc. Pacif.* **83**, 95–97 (1971).
- Mundt, R. & Bastian, U. UVB photometry of young emission-line objects. *Astron. Astrophys. Suppl.* **39**, 245–250 (1980).
- Herbig, G. H. & Bell, K. R. Third catalog of emission-line stars of the Orion population. *Lick Observ. Bull.* No. 1111 (1988).
- Graham, J. A. Objects associated with low-mass star formation in the Gum Nebula. *Astrophys. J.* **302**, 352–362 (1986).
- Hirth, G. A., Mundt, R. & Solf, J. Spatial and kinematic properties of the forbidden emission line region of T Tauri stars. *Astron. Astrophys. Suppl.* **126**, 437–469 (1997).
- Solf, J. in *ESO Workshop on Low Mass Star Formation and Pre-Main Sequence Objects* (ed. Reipurth, B.) 399–406 (ESO, Garching, 1989).
- Gorti, U. & Bhatt, H. C. Orbital decay of protostellar binaries in molecular clouds. *Mon. Not. R. Astron. Soc.* **283**, 566–576 (1996).
- Clarke, C. J. & Syer, D. Low-mass companions to T Tauri stars: a mechanism for rapid-rise FU Orionis outbursts. *Mon. Not. R. Astron. Soc.* **278**, L23–L27 (1996).
- Herbig, G. H. The young cluster IC 348. *Astrophys. J.* **497**, 736–758 (1998).
- Reipurth, B., Bally, J., Graham, J. A., Lane, A. P. & Zealey, W. The jet and energy source of HH 34. *Astron. Astrophys.* **164**, 51–66 (1986).
- Reipurth, B. & Heathcote, S. The jet and energy source of HH 46/47. *Astron. Astrophys.* **246**, 511–534 (1991).
- Reipurth, B. The HH 111 jet and multiple outflow episodes from young stars. *Nature* **340**, 42–45 (1989).
- Gredel, R. & Reipurth, B. An infrared counter-flow in the HH 111 jet complex. *Astron. Astrophys.* **289**, L19–L22 (1994).
- Zinnecker, H., McCaughrean, M. J. & Rayner, J. T. A symmetrically pulsed jet of gas from an invisible protostar in Orion. *Nature* **394**, 862–865 (1998).
- Hirth, G. A., Mundt, R., Solf, J. & Ray, T. P. Asymmetries in bipolar jets from young stars. *Astrophys. J.* **427**, L99–L102 (1994).

Acknowledgements. We thank A. Raga and D. Hollenbach for discussions. L.-Å. Nyman for a CO measurement, D. Theil for help in producing the figures, and M. McCaughrean and R. Ouyed for comments and advice.

Correspondence and requests for materials should be addressed to B.R. (e-mail address: reipurth@casa.colorado.edu).

Spin domains in ground-state Bose-Einstein condensates

J. Stenger, S. Inouye, D. M. Stamper-Kurn, H.-J. Miesner, A. P. Chikkatur & W. Ketterle

Department of Physics and Research Laboratory of Electronics, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA

Bose-Einstein condensates—a low-temperature form of matter in which a macroscopic population of bosons occupies the quantum-mechanical ground state—have been demonstrated for weakly interacting, dilute gases of alkali-metal^{1–3} and hydrogen²⁵ atoms. Magnetic traps are usually employed to confine the condensates, but have the drawback that spin flips in the atoms lead to untrapped states. For this reason, the spin orientation of the trapped alkali atoms cannot be regarded as a degree of freedom. Such condensates are therefore described by a scalar order parameter, like the spinless superfluid ⁴He. In contrast, a recently realized optical trap⁴ for sodium condensates confines atoms independently of their spin orientations. This offers the possibility of studying ‘spinor’ condensates in which spin comprises a degree of freedom, so that the order parameter is a vector rather than scalar quantity. Here we report the observation of equilibrium states of sodium spinor condensates in an optical trap. The freedom of spin orientation leads to the formation of spin domains in an external magnetic field, which can be either miscible or immiscible with one another.

A variety of new phenomena are predicted^{5–7} for spinor condensates, such as spin textures, propagation of spin waves and coupling between superfluid flow and atomic spin. To date, such effects could only be studied in superfluid ³He, which can be

described by Bose–Einstein condensation of Cooper pairs of quasiparticles having both spin and orbital angular momentum⁸. Compared to the strongly interacting ³He, the properties of weakly interacting Bose–Einstein condensates of alkali-metal gases can be calculated by mean field theories in a much more straightforward and simple way.

Other systems which go beyond the description with a single scalar order parameter are condensates of two different hyperfine states of ⁸⁷Rb confined in magnetic traps. Recent experimental studies have explored the spatial separation of the two components^{9,10} and their relative phase¹¹. Several theoretical papers describe their structure^{12–18} and their collective excitations^{19–22}.

Compared to these two-component condensates, spinor condensates have several new features, including the vector character of the order parameter and the changed role of spin relaxation collisions which allow for population exchange among hyperfine states without trap loss. In contrast, for ⁸⁷Rb experiments trap loss due to spin relaxation severely limits the lifetime.

Here we consider an $F = 1$ spinor condensate subject to spin relaxation, in which two $m_F = 0$ atoms can collide and produce an $m_F = +1$ and an $m_F = -1$ atom and vice versa. (Here, F denotes the angular momentum per atom and m_F its magnetic quantum number.) We investigate the distribution of hyperfine states and the spatial distribution in equilibrium assuming conservation of the total spin.

The ground state spinor wavefunction is found by minimizing the free energy⁵;

$$K = \int d^3r n \left[V + \frac{c_0 n}{2} + \frac{c_2 n}{2} \langle \mathbf{F} \rangle^2 + E_{ze} - p_0 \langle F_z \rangle \right] \quad (1)$$

where kinetic energy terms are neglected in the Thomas–Fermi approximation which is valid as long as the dimension of spin domains (typically 50 μm) is larger than the penetration depth¹⁸ (typically 1 μm). V is trapping potential, n is the density, r is the spatial coordinate and E_{ze} is the Zeeman energy in an external magnetic field. The Lagrange multiplier p_0 accounts for the total spin conservation. The mean field energy in equation (1) consists of a spin-independent part proportional to c_0 and a spin-dependent part proportional to $c_2 \langle \mathbf{F} \rangle^2$. The coefficients c_0 and c_2 are related to the scattering lengths a_0 and a_2 for two colliding atoms with total angular momentum $F_{\text{tot}} = 0$ or $F_{\text{tot}} = 2$ by $c_0 = 4\pi\hbar^2 \bar{a}/M$ and $c_2 = 4\pi\hbar^2 \Delta a/M$ with $\bar{a} = (2a_2 + a_0)/3$, $\Delta a = (a_2 - a_0)/3$, and M for the atomic mass (ref. 5). The spin-dependent interaction originates from the term $c_2 \mathbf{F}_1 \cdot \mathbf{F}_2$ in the interaction of two atoms, which is ferromagnetic for $c_2 < 0$ and antiferromagnetic for $c_2 > 0$.

In the Bogoliubov approach, the many-body ground-state wavefunction is represented by the spinor wavefunction;

$$\Psi(\mathbf{r}) = \sqrt{n(\mathbf{r})} \zeta(\mathbf{r}) = \sqrt{n(\mathbf{r})} (\zeta_+(\mathbf{r}), \zeta_0(\mathbf{r}), \zeta_-(\mathbf{r})) \quad (2)$$

where ζ_+ , ζ_0 , ζ_- denote the amplitudes for the $m_F = +1, 0, -1$ states, respectively, and $|\zeta|^2 = 1$.

The Zeeman energy E_{ze} is given by;

$$E_{ze} = E_+ |\zeta_+|^2 + E_0 |\zeta_0|^2 + E_- |\zeta_-|^2 = E_0 - \tilde{p} \langle F_z \rangle + q \langle F_z^2 \rangle \quad (3)$$

where E_+ , E_0 , E_- are the Zeeman energies of the $m_F = +1, 0, -1$ states, $2q \equiv E_+ + E_- - 2E_0$ is the Zeeman energy difference in a spinflip collision, and $2\tilde{p} \equiv E_- - E_+$. The E_0 term can be included in the trapping potential V . The parameter $\tilde{p}$ can be combined with the Lagrange multiplier p_0 to give $p \equiv \tilde{p} + p_0$.

In the following we determine the spinor which minimizes the spin-dependent part K_s of the free energy:

$$K_s = c \langle \mathbf{F} \rangle^2 - p \langle F_z \rangle + q \langle F_z^2 \rangle \quad (4)$$

where $c = c_2 n/2$. The minimization of equation (4) for different values of the parameters c , p and q is straightforward, and is shown

graphically in the form of spin-domain diagrams in Fig. 1.

Experimentally, the values of c , p and q can be varied arbitrarily, representing any region of the spin-domain diagram. The magnitude (but not the sign) of the coefficient c is varied by changing the density n , either by changing the trapping potential, or by studying condensates with different numbers of atoms. In this study, the axial length of the trapped condensate is more than 60 times larger than its radial size, and thus we consider the system one-dimensional, and integrate over the radial coordinates, obtaining $n = 2n_0/3$ where n_0 is the density at the radial centre. This integration assumes a parabolic density profile within the Thomas–Fermi approximation. The value of q can be changed by applying a weak external bias field B_0 : q then corresponds to the quadratic Zeeman shift which is proportional to B_0^2 . The coefficient p arises both from the linear Zeeman shift and from the Lagrange multiplier p_0 which is determined by the total spin of the system. For a system with zero total spin in a homogenous bias field B_0 , p_0 cancels the linear Zeeman shift due to B_0 , yielding $p = 0$. Positive (negative) values of p are achieved for condensates with a positive (negative) overall spin. Finally, the coefficients can be made to vary spatially across the condensate. In particular, applying a field gradient B' along the axis of the trapped condensate causes p to vary along the condensate length. For a condensate with zero total spin, p is proportional to $B'z$ where z is the axial coordinate with $z = 0$ at the centre of the condensate. Thus, the condensate samples a vertical line in the spin-domain diagrams of Fig. 1. The centre of this line lies at $p = 0$, and its length is given by the condensate length scaled by B' .

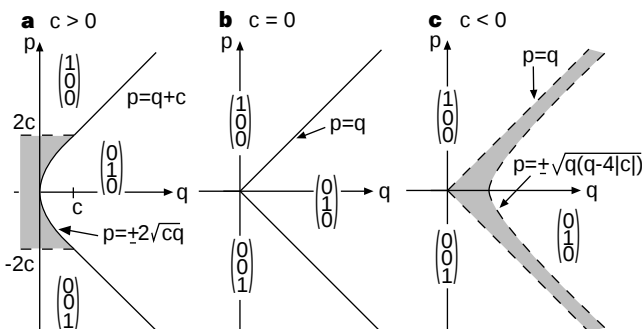


Figure 1 Spin-domain diagrams for condensates with $F = 1$. The structure of the ground-state spinor is shown as a function of the linear ($\sim p$) and quadratic ($\sim q$) Zeeman energies. Hyperfine components are mixed inside the shaded regions. Solid lines indicate a discontinuous change of state populations whereas dashed lines indicate a gradual change. The behaviour for $q < 0$ is also shown, although it is not relevant for the experiment reported here. For $c = 0$, the Zeeman energy causes the cloud to separate into three domains with $m_F = +1, 0, -1$ and with boundaries at $|p| = q$, as shown in **b**. For $c_2 \neq 0$, the mean field energy shifts the boundary region between domains and leads to regions of overlapping spin components. In the antiferromagnetic case (**a**), the $m_F = 0$ component and the $m_F = \pm 1$ components are immiscible (including the kinetic energy terms in equation (1) would lead to a thin boundary layer) and the boundary occurs at $|p| = q + c$. For small bias fields, with $q < c$ and $|p| < 2c$, the $m_F = 0$ domain is bordered by domains in which $m_F = \pm 1$ components are mixed. The ratio of the $m_F = \pm 1$ populations in these regions does not depend on q , but is given by $|\zeta_+|^2/|\zeta_-|^2 = (2c + p)/(2c - p)$. In this region of small fields, the boundary to the $m_F = 0$ component lies at $|p| = 2\sqrt{cq}$. In the ferromagnetic case (**c**) all three components are generally miscible, and have no sharp boundaries. Pure $m_F = 0$ domains occur for $|p| \leq \sqrt{q(q - 4|c|)}$ and pure $m_F = \pm 1$ domains for $|p| > q$. Here, in contrast to the antiferromagnetic case, a pure $m_F = 0$ condensate is skirted by regions where it is mixed predominantly with either the $m_F = -1$ or the $m_F = +1$ component. The contribution of the third component is very small ($< 2\%$). In all mixed regions, the $m_F = 0$ component is never the least populated of the three spin components. This qualitative feature can be used to rule out the possibility that $F = 1$ sodium atoms have ferromagnetic interactions.

Our experimental study of spinor condensates required techniques to selectively prepare and probe condensates in arbitrary hyperfine states. Spinor condensates were prepared in several steps. Laser cooling and evaporative cooling were used to produce sodium condensates in the $m_F = -1$ state in a cloverleaf magnetic trap²³. The condensates were then transferred into an optical dipole trap consisting of a single focused infrared laser beam⁴. After the spin preparation, a bias field B_0 and a field gradient B' were applied for a variable amount of time (as long as 30 s), during which the atoms relaxed towards their equilibrium distribution (Fig. 2).

The profiles in Fig. 3 were obtained from vertical cuts through absorption images. They provide clear evidence of antiferromagnetic interaction. (The opposite case is predicted for the $^{87}\text{Rb } F = 1$ spin multiplet²⁴.) The spin structure is consistent with the corresponding spin-domain diagram in Fig. 1a. Overlapping $m_F = \pm 1$ clouds as observed are incompatible with the assumption of ferromagnetic interaction.

The strength $c = (50 \pm 20)$ Hz of the antiferromagnetic interaction was estimated by determining z_b , the location of the $m_F = 0$ to the $m_F = \pm 1$ boundary, and by plotting $p = \mu B' z_b$ versus the quadratic Zeeman shift $q = \hat{q} B_0^2$ (Fig. 4). The constants μ and $\hat{q}$ are defined in the figure legend. With $n = (2.9 \pm 0.5) \times 10^{14} \text{ cm}^{-3}$, the difference between the scattering

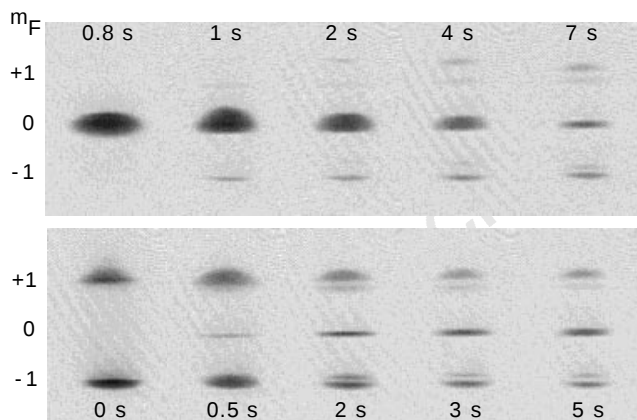


Figure 2 Formation of ground-state spin domains. Absorption images of ballistically expanding spinor condensates show both the spatial and hyperfine distributions. Arbitrary populations of the three hyperfine states were prepared using r.f. transitions (Landau-Zener sweeps)⁴. At a bias field of ~ 40 G, the transitions from $m_F = -1$ to $m_F = 0$ and from $m_F = 0$ to $m_F = +1$ differ in frequency by ~ 0.9 MHz due to the quadratic Zeeman shift, and they could be driven separately. The images of clouds with various dwell times in the trap show the evolution to the same equilibrium for condensates prepared in either a pure $m_F = 0$ state (upper row) or in equally populated $m_F = \pm 1$ states (lower row). Between 5 and 15 s dwell time, the distribution did not change significantly, although the density decreased due to three-body recombination. The bias field during the dwell time was $B_0 = 20$ mG, and the field gradient was $B' = 11 \text{ mG cm}^{-1}$. These images were taken after the optical trap was suddenly switched off, and the atoms were allowed to expand. Due to the large aspect ratio (typically 60), the expansion was almost purely in the radial directions. All the mean-field energy was released after < 1 ms, after which the atoms expanded as free particles. Thus, a magnetic field gradient, which was applied after 5 ms time of flight to yield a Stern-Gerlach separation of the cloud, merely translated the three spin components without affecting their shapes. In this manner, the single time-of-flight images provided both a spatial and spin-state description of the trapped cloud. Indeed, the shapes of the three clouds fit together to form a smooth total density distribution. After a total time of flight of 25 ms, the atoms were optically pumped into the $F = 2$ hyperfine state and observed using the $m_F = +2$ to $m_F = +3$ cycling transition. This technique assured the same transition strength for atoms originating from different spin states. The size of the field of view for a single spinor condensate is 1.7×2.7 mm.

lengths can be determined to $a_2 - a_0 = 3\Delta a = (3.5 \pm 1.5)a_B = (0.19 \pm 0.08) \text{ nm}$ where a_B is the Bohr radius. This result is in rough agreement with a theoretical calculation of $a_2 - a_0 = (5.5 \pm 0.5)a_B$ (ref. 24). The antiferromagnetic interaction energy corresponds to 2.5 nK in our condensates. Still, the magnetostatic (ferromagnetic) interaction between the atomic magnetic moments is about ten times weaker. We note that the optically trapped samples in which the domains were observed were at a temperature of the order of 100 nK, far larger than the antiferromagnetic energy. The formation of spin domains occurs only in a Bose-Einstein condensate.

Figure 3c shows a profile of the density distribution for a cloud at $B_0 = 20$ mG and almost-cancelled gradient ($B' < 2 \text{ mG cm}^{-1}$). No $m_F = 0$ region can be identified. The cloud was prepared with a small total angular momentum. Due to the almost-zero gradient and the non-zero angular momentum, the cloud corresponds to a point in the shaded region in Fig. 1a, rather than a vertical line with no offset as discussed before with finite gradients and zero angular momentum. The different widths of the profiles are probably caused by residual field inhomogeneities. Figure 3c demonstrates the complete miscibility of the $m_F = \pm 1$ components.

For a homogenous two-component system the criterion for miscibility (immiscibility) is^{14,15,18} $a_{ab} < (>) \sqrt{a_a a_b}$, when the mean

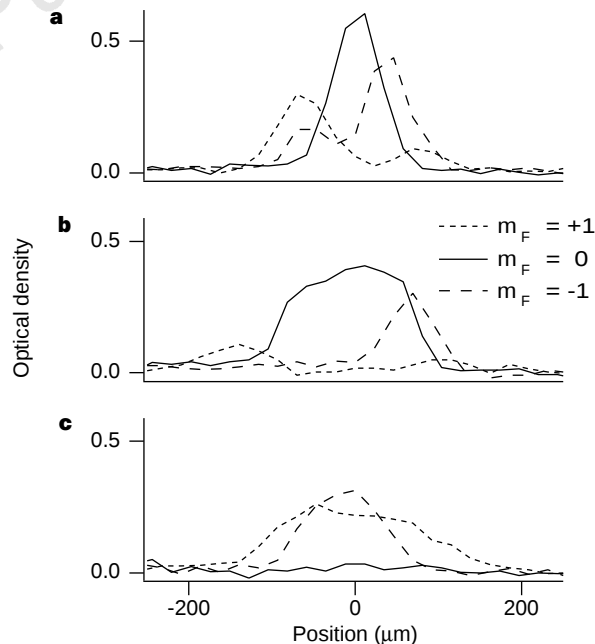


Figure 3 Miscible and immiscible spin domains. Axial column density profiles of spinor Bose-Einstein condensates are shown, obtained from time-of-flight absorption images as in Fig. 2. The profiles of the $m_F = \pm 1$ components were shifted to undo the Stern-Gerlach separation. At low bias fields (**a**), the $m_F = 0$ component was skirted on both sides by $m_F = \pm 1$ components with significant $m_F = \mp 1$ admixtures, thus demonstrating the antiferromagnetic interaction (also visible in Fig. 2). At higher fields (**b**), the $m_F = \pm 1$ components are pushed apart further by a larger $m_F = 0$ component, and the $m_F = \mp 1$ admixtures vanish. They could not be resolved for quadratic Zeeman energies $q > 20$ Hz. The antiferromagnetic interaction leads to immiscibility of the $m_F = 0$ and the $m_F = \pm 1$ components. The kinetic energy in this boundary region, which is small compared to the total mean-field energy, is released in the axial direction. Due to this axial expansion of the cloud in the time of flight, and due to imperfections in the imaging system including the limited pixel resolution, the $m_F = 0$ to $m_F = \pm 1$ boundary is not sharp. Panel **c** demonstrates the complete miscibility of the $m_F = \pm 1$ components. The magnetic field parameters were $B_0 = 20$ mG, $B' = 11 \text{ mG cm}^{-1}$ in **a**, $B_0 = 100$ mG, $B' = 11 \text{ mG cm}^{-1}$ in **b**, and $B_0 = 20$ mG, $B' < 2 \text{ mG cm}^{-1}$ in **c**.

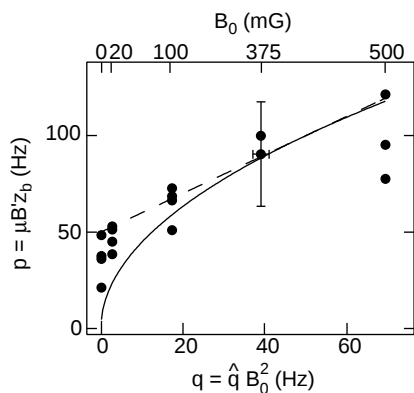


Figure 4 Estimate of the antiferromagnetic interaction energy c . Plotted is the linear Zeeman energy $p = |\mu B' z_b|$ at the boundary between the $m_F = 0$ and $m_F = \pm 1$ regions versus the quadratic Zeeman shift $q = \hat{q} B_0^2$. Here $\hat{q} = g_s^2 \mu_B^2 / 16 \hbar^2 \nu_{\text{HFS}} = 278 \text{ Hz G}^{-2}$, and $\mu = g_s \mu_B / 4 \hbar = 700 \text{ kHz G}^{-1}$, g_s denotes the electron g -factor, ν_{HFS} the hyperfine splitting frequency, and μ_B the Bohr magneton. The solid line is a fit of the function $|p| = 2\sqrt{qc}$ for $q < c$ and $|p| = q + c$ for $q > c$. Extrapolating the linear part to zero bias field (dashed line) yields $c = (50 \pm 20) \text{ Hz}$. The data points at a given bias field represent $p = \mu B' z_b$ for different gradient fields B' and thus $m_F = 0$ regions of different size. The scatter of these points is mainly due to residual magnetic field inhomogeneities, resulting in small deviations of the local gradient B' . The error bar represents the relative error of all data points of 30% in p and 5% in q , as estimated from the uncertainties in the magnetic-field calibration. Furthermore, the limited pixel resolution and contributions of the kinetic energy in the condensate to the axial expansion enhance the errors for the determination of z_b of small $m_F = 0$ regions.

field energy is parametrized as $(2\pi\hbar^2/M)(n_a^2 a_a + n_b^2 a_b + 2n_a n_b a_{ab})$. Here, $n_{a,b}$ and $a_{a,b}$ are densities and scattering lengths for the components a and b , and the scattering length a_{ab} characterizes the interactions between particles a and b . In our spinor condensate with mixtures of the $m_F = \pm 1$ components, we have $a_{-1} = a_{+1} = \bar{a} + \Delta a$ and $a_{-1+1} = \bar{a} - \Delta a$. Thus $\Delta a > 0$, as experimentally observed, implies miscibility. For a mixture of the $m_F = 1$ and $m_F = 0$ components, we find $a_0 = \bar{a}$, $a_{+1} = \bar{a} + \Delta a$ and $a_{0+1} = \bar{a} + \Delta a$, corresponding to immiscibility. For the ^{87}Rb experiments^{9,10}, it is not clear whether the two components are miscible or overlap only in a surface region due to kinetic energy (E. A. Cornell, personal communication).

All regions in the spin-domain diagrams are accessible with our experimental technique, and thus any combination of the three hyperfine components can be realized by applying small external magnetic fields. Of special interest for future work is the zero-magnetic-field case, where the rotational symmetry should be spontaneously broken. We observed both miscibility and immiscibility of hyperfine components. Thus the dynamics and possible metastable configurations⁷ of two interpenetrating, miscible superfluid components ($m_F = \pm 1$) with arbitrary admixtures of an immiscible component ($m_F = 0$) can now be studied. □

Received 18 June; accepted 3 September 1998.

1. Anderson, M. H., Ensher, J. R., Matthews, M. R., Wieman, C. E. & Cornell, E. A. Observation of Bose-Einstein condensation in a dilute atomic vapor. *Science* **269**, 198–201 (1995).
2. Davis, K. B. *et al.* Bose-Einstein condensation in a gas of sodium atoms. *Phys. Rev. Lett.* **75**, 3969–3973 (1995).
3. Bradley, C. C., Sackett, C. A. & Hulet, R. G. Bose-Einstein condensation of lithium: Observation of limited condensate number. *Phys. Rev. Lett.* **78**, 985–989 (1997).
4. Stamper-Kurn, D. M. *et al.* Optical confinement of a Bose-Einstein condensate. *Phys. Rev. Lett.* **80**, 2027–2030 (1998).
5. Ho, T.-L. Spinor Bose condensates in optical traps. *Phys. Rev. Lett.* **81**, 742–745 (1998).
6. Ohmi, T. & Machida, K. Bose-Einstein condensation with internal degrees of freedom in alkali atom gases. *J. Phys. Soc. Jpn* **67**, 1822–1825 (1998).
7. Law, C. K., Pu, H. & Bigelow, N. P. Quantum spins mixing in spinor Bose-Einstein condensates. *Phys. Rev. Lett.* (submitted); preprint cond-mat/9807258 at (<http://xxx.lanl.gov>) (1998).
8. Vollhardt, D. & Wölfle, P. *The Superfluid Phases of ^4He* (Taylor-Francis, London, 1990).
9. Myatt, C. J., Burt, E. A., Ghrist, R. W., Cornell, E. A. & Wieman, C. E. Production of two overlapping Bose-Einstein condensates by sympathetic cooling. *Phys. Rev. Lett.* **78**, 586–589 (1997).

10. Hall, D. S., Matthews, M. R., Ensher, J. R., Wieman, C. E. & Cornell, E. A. The dynamics of component separation in a binary mixture of Bose-Einstein condensates. *Phys. Rev. Lett.* **81**, 1539–1542 (1998).
11. Hall, D. S., Matthews, M. R., Wieman, C. E. & Cornell, E. A. Measurements of relative phase in binary mixtures of Bose-Einstein condensates. *Phys. Rev. Lett.* **81**, 1543–1546 (1998).
12. Siggia, E. D. & Ruckenstein, A. E. Bose condensation in spin-polarized atomic hydrogen. *Phys. Rev. Lett.* **44**, 1423–1426 (1980).
13. Ho, T.-L. & Shenoy, V. B. Binary mixtures of Bose condensates of alkali atoms. *Phys. Rev. Lett.* **77**, 3276–3279 (1996).
14. Timmermans, E. Phase separation in Bose-Einstein condensates. *Phys. Rev. Lett.* (submitted); preprint cond-mat/9709301 at (<http://xxx.lanl.gov>) (1997).
15. Esry, B. D., Greene, C. H., Bruke, J. P. & Bohn, J. L. Hartree-Fock theory for double condensates. *Phys. Rev. Lett.* **78**, 3594–3597 (1997).
16. Öhberg, P. & Stenholm, S. Hartree-Fock treatment of the two-component Bose-Einstein condensate. *Phys. Rev. A* **57**, 1272–1279 (1998).
17. Pu, H. & Bigelow, N. P. Properties of two-species Bose condensates. *Phys. Rev. Lett.* **80**, 1130–1133 (1998).
18. Ao, P. & Chui, S. T. Binary Bose-Einstein condensate mixtures in weakly and strongly segregated phases. *Phys. Rev. A* **58**(6) (in the press).
19. Busch, T., Cirac, J. I., Pérez-García, V. M. & Zoller, P. Stability and collective excitations of a two-component Bose-Einstein condensed gas: a moment approach. *Phys. Rev. A* **56**, 2978–2983 (1997).
20. Graham, R. & Walls, D. Collective excitations of trapped binary mixtures of Bose-Einstein condensed gases. *Phys. Rev. A* **57**, 484–487 (1998).
21. Pu, H. & Bigelow, N. P. Collective excitations, metastability, and nonlinear response of a trapped two-species Bose-Einstein condensate. *Phys. Rev. Lett.* **80**, 1134–1137 (1998).
22. Esry, B. D. & Greene, C. H. Low-lying excitations of double Bose-Einstein condensates. *Phys. Rev. A* **57**, 1265–1271 (1998).
23. Mewes, M.-O. *et al.* Bose-Einstein condensation in a tightly confining dc magnetic trap. *Phys. Rev. Lett.* **77**, 416–419 (1996).
24. Burke, J. P., Greene, C. H. & Bohn, J. L. Multichannel cold collisions: simple dependencies on energy and magnetic field. *Phys. Rev. Lett.* **81**, 3355–3358 (1998).
25. Fried, D. G. *et al.* Bose-Einstein condensation of atomic hydrogen. *Phys. Rev. Lett.* (in the press).

Acknowledgements. We acknowledge discussions with J. Ho and C. Greene. This work was supported by the Office of Naval Research, NSF, Joint Services Electronics Program (ARO), NASA, and the David and Lucile Packard Foundation. J.S. acknowledges support from the Alexander von Humboldt foundation, D.M.S.-K. from the JSEP Graduate Fellowship Program, and A.P.C. from the NSF.

Correspondence and requests for materials should be addressed to J.S. (e-mail: stenger@amo.mit.edu).

Observation of smectic and moving-Bragg-glass phases in flowing vortex lattices

F. Pardo*, F. de la Cruz*, P. L. Gammel†, E. Bucher† & D. J. Bishop†

* Centro Atómico Bariloche and Instituto Balseiro, Comisión Nacional de Energía Atómica, 8400 S. C. de Bariloche, R. N., Argentina

† Bell Laboratories, Lucent Technologies, Murray Hill, New Jersey 07974, USA

The defining characteristic of the superconducting state is its ability to carry electrical currents without loss. The process by which it does this has been extensively studied for decades but there are still many unresolved issues. In particular, the critical current, which is the maximum electrical current that a superconductor can carry without loss, remains a poorly understood concept at the microscopic level. In a type II superconductor, a flux-line lattice (FLL) forms if a magnetic field between H_{c1} and H_{c2} , the lower and upper critical fields, is applied: flowing electrical currents will exert a force on this FLL. If the FLL remains pinned, the current flows without loss of energy and the effective resistance remains zero. However, if the lattice moves in response to the current, energy is dissipated and the zero-resistance state is lost. Because of its relevance to the critical current, the types of structures that these moving lattices can form have attracted much recent theoretical attention^{1–4}. Here we report magnetic decoration studies of flowing vortex lattices which show evidence for a transition, as a function of increasing flux density, from a layered (or smectic) FLL² to a more well-ordered moving Bragg glass¹.

Previous work on this problem has uncovered a number of results on the structures that can be formed by flowing vortices. Early imaging studies⁵ using small-angle neutron scattering (SANS) confirmed less-direct evidence from transport studies⁶, and found that while pinned lattices are disordered, flowing lattices can be

quite well-ordered with resolution-limited diffraction peaks. An important theoretical advance was made on the basis of numerical simulations by Koshelev and Vinokur⁴ who argued that there could be three distinct regimes of flow. For applied currents below the critical current (J_c), these authors found a pinned, moderately disordered solid; at J_c they saw a regime of quite disordered plastic flow; and for currents well above J_c , they found a regime where the moving lattice was quite well ordered, perhaps crystalline. This prediction of a plastic flow regime was confirmed by both SANS⁷ and magnetic decoration studies^{8,9} of the flowing FLL at J_c .

More recently, Giarmarchi and Le Doussal¹ have argued for the existence of a moving-Bragg-glass phase for currents well above J_c . This is a phase free from topological defects, with power-law decay of the positional correlations which are anisotropic with respect to the flow direction. Balents, Marchetti and Radzihovsky² have argued instead for a smectic phase. Such a phase is layered in the sense that the vortices flow in channels and are well correlated perpendicular to their flow, but uncorrelated along it. This is consistent with recent numerical simulations³.

In the magnetic decoration technique used here, the FLL is visualized by evaporating $\sim 50\text{-\AA}$ magnetic particles onto the surface of a superconductor held below its transition temperature T_c with a magnetic field applied. The particles follow the field lines of the vortices at the surface of the sample, and land where the vortices are located. Because each pile of magnetic particles has a finite size, the technique is limited to fields below several hundred oersteds. Fortunately, this range of fields covers an interesting region where we can 'tune' the vortex interaction strength relative to the pinning strength, and see a crossover from a smectic structure in the low-field, disorder-dominated limit to a moving Bragg glass in the high-field, interaction-dominated limit.

Our samples were high-quality, single crystals of NbSe₂ with typical dimensions of $0.5 \times 0.5 \times 0.2$ mm. They were cleaved just before use to assure a clean surface for the measurement. Samples were cooled down to 4.2 K in a magnetic field applied in the c direction. At low temperatures, the field was removed and the sample was decorated after a certain period of time (t).

Figure 1 shows data from a SQUID (superconducting interference device) measurement of the magnetization of our sample taken as a function of time (t) after the 36 Oe field was removed. The flow velocity (v) at the edge of the sample was extracted from the magnetization using the relation $v = [\partial m(t)/\partial t][A/m(t)P]$ where A is the area, P the perimeter, and $m(t)$ the magnetization. The average

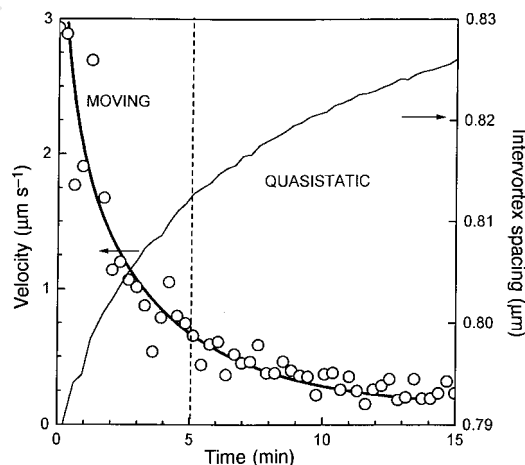


Figure 1 Average vortex velocity and intervortex spacing a_0 as a function of time during flux creep for a sample of NbSe₂ at 4.2 K. A field of 35 Oe was originally applied and then removed at $t = 0$. In the moving regime, the vortices move more than the intervortex spacing during an experiment; in the quasistatic regime, the vortices move less than an intervortex spacing during an experiment.

lattice constant plotted was obtained using $a_0(t) = [\Phi_0 V/m(t)]^{1/2}$ where V is the sample volume and Φ_0 is the flux quantum. With a one-second decoration time, the plot of the vortex velocity defines two regimes. If the velocity is greater than one vortex lattice constant per second, we find images of vortices in motion. For the regime of vortex velocities less than one lattice constant per second, we obtain quasistatic images of slowly moving vortices. We can work in both regimes. Because of the critical-state profile (independent of velocity regime), we always find a quasistatic, homogeneous structure in the centre of the sample and the critical-state region at the edge. The data shown here is always from the critical-state region.

Our imaging data are shown in Fig. 2. There are three types of images shown for four different regimes (Fig. 2a–d). The first column shows the real-space images (RS) of the FLL, the second the Fourier-transformed (FT) data and the third a Fourier-filtered (FF) real-space image. In the first column, the flow direction is indicated with an arrow. The row of images in Fig. 2a is low-field/low-velocity data, Fig. 2b is low-field/high-velocity, Fig. 2c is high-field/low-velocity and Fig. 2d is high-field/high-velocity data.

Because the data in row Fig. 2a are in the quasistatic regime, the

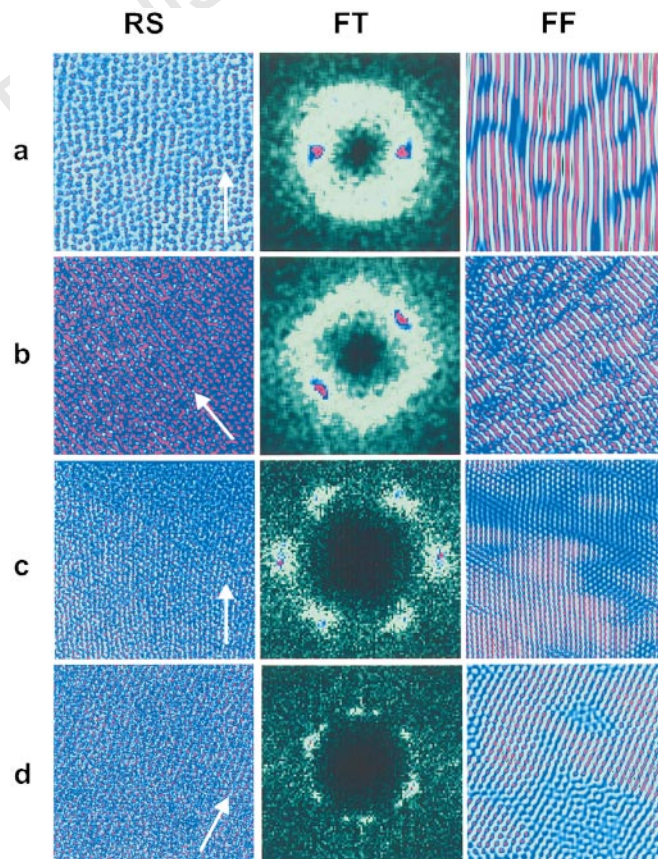


Figure 2 Images of the flux line lattice at different fields and flow velocities. Shown are real space (RS; first column) Fourier-transform space (FT; second column) and Fourier-filtered (FF; third column) images. The real-space images were digitized, flat-fielded on a 20×20 kernel and Lee filtered with a 5×5 kernel. To obtain the Fourier-filtered images, the real and imaginary components of each pixel in the Fourier-transform image were multiplied by a smooth sigmoidal function of their magnitude with a threshold of $\sim 1/3$ of the peak magnitude of the FT over the entire image. These data were then inverse-transformed and the absolute value of the inverse transform was used to construct the third column. **a**, Low-field, low-velocity data (5 Oe, $0.2 \mu\text{m s}^{-1}$); **b**, low-field, high-velocity data (5 Oe, $2.5 \mu\text{m s}^{-1}$); **c**, high-field, low-velocity data (20 Oe, $0.2 \mu\text{m s}^{-1}$); **d**, high-field, high-velocity data (25 Oe, $3 \mu\text{m s}^{-1}$). Rows **a** and **b** show a smectic vortex lattice; **c** and **d** show a moving-Bragg-glass state. The arrows in the first column show the direction of flow.

RS image resolves individual vortices. We find rows of vortices aligned along the flow direction and not an isotropic, hexagonal pattern as would be seen for a static lattice. The FF image has the individual vortices filtered out, but the row or smectic structure in the data remains; one can clearly see the rows which make up the domains of flow. The FT data shows the signature of a smectic phase. The two sharp peaks in the FT image show that the channels are well-oriented transverse to the flow of vortices, and the broad lines of scattering show that the vortices are liquid-like in the channels with short-range positional correlations along the direction of flow.

The data in Fig. 2b are no longer in the quasistatic regime, and the RS image no longer resolves individual vortices as clearly but the same smectic pattern can be seen. A careful look at this image reveals the existence of decoupled motion between channels: there are blurred and static channels. In fact, the FF image suggests that the domains of flow are smaller at this velocity than at a smaller velocity.

A quantitative analysis of the data in Fig. 2a and b gives the FT radial width along the flow to be 35% and 20% wider for the low- and high-velocity cases, respectively, than the width transverse to the flow. This structure is quite similar to smectic liquid crystals in which there is a well-defined layer structure with the molecules in each layer liquid-like. This structure agrees in detail with that predicted by Balents, Marchetti and Radzihovsky² who argued that in this regime the FT should show two sharp peaks for the transverse order, and only diffuse peaks or two lines of scattering for the other four Bragg peaks.

Quite different behaviour is seen in Fig. 2c; this image shows six-fold order, suggesting that the correlation lengths in both directions are finite. The remaining topological defects (a few per cent in density) arise mainly because of the small field gradient. The FT image confirms this, as one sees six well-defined peaks although there is still a one-dimensional modulation with two of the peaks brighter than the other four, indicating that the positional correlation length perpendicular to the flow direction is still longer than that parallel to the flow. These data suggest that the increased field allows the vortex–vortex interactions to dominate over the disorder, and is the equivalent to reducing the disorder. The FT data is similar to what one would find for a tilted hexatic liquid crystal where there is long-ranged orientational order with a one-dimensional modulation superimposed due to the tilt of the liquid-crystal molecule^{10,11}. It is quite similar to the predicted moving-Bragg-glass phase and to that found in numerical simulations in three dimensions¹².

For the data in row Fig. 2d the RS image does not resolve individual vortices as the system is not in the quasistatic regime. Such a pattern is similar to that found in ref. 13. The FT image shows six well-resolved peaks indicative of an hexagonal lattice, again with an obvious one-dimensional modulation superimposed with two of the peaks brighter than the other four. The FF filtered image shows coherent motion evidenced by large flow domains.

The data in Fig. 2c can be quantitatively analysed, and the various correlation functions extracted. The lattice is hexatic, with the orientational correlation longer-ranged than the positional correlation. This is because in the raw FT data there are six sharp peaks, which indicates that the orientational correlation extends over the entire field of view of the data set while the positional correlation only extends over several lattice constants. More interesting is the fact that both the positional and orientational correlation functions are anisotropic with respect to the flow direction. For the low-velocity, high-field data the radial widths of the peaks are 4% wider, and the azimuthal widths are 12% wider, for the direction along the flow direction. This indicates that both the positional and orientational correlation lengths are longer perpendicular to the flow direction than parallel to it, as would be expected for a moving-Bragg-glass phase of a flowing vortex lattice. The low value of the anisotropy found would be consistent with the low-velocity regime in the theory¹. For the high-field, high-velocity data, the effect is larger,

with the radial widths differing by 30% and the azimuthal widths by 75% between the directions parallel and perpendicular to the flow direction; widths parallel to the motion are the largest. But this increase could be partially due to the ‘decoration in motion’ process that selectively diffuses positions along the velocity direction.

We note that there are two different types of states in the literature which have been given the name moving Bragg glass. In the original proposal by Giarmarchi-Le Doussal¹, the correlation lengths for this state were expected to be anisotropic as found here but longer in the flow direction than perpendicular to it. This is in contrast to our data where the correlation lengths perpendicular to the flow direction are longer. In more recent work by Balents, Marchetti and Radzihovsky², a Bragg glass phase for flowing vortices was found but one in which anisotropy of the positional correlation agrees with our results.

Our data show some dynamics of the moving vortex state at and around the critical current of a type II superconductor. As predicted by Koshelev and Vinokur⁴, flow can, contrary to one’s intuition, induce order in the FLL of a type II superconductor. Our results confirm this notion but also hint at how subtle and complex that order can be. During the preparation of this Letter, theoretical work was published that addressed several aspects of the vortex dynamic phases^{14–16}. □

Received 17 September 1997; accepted 20 August 1998.

1. Giarmarchi, T. & Le Doussal, P. Moving glass phase of driven lattices. *Phys. Rev. Lett.* **76**, 3408–3411 (1996).
2. Balents, L., Marchetti, C. M. & Radzihovsky, L. Comment on moving glass phase of driven lattices. *Phys. Rev. Lett.* **78**, 751 (1997).
3. Moon, K., Scalettar, T. & Gergely, T. Dynamical phases of driven vortex systems. *Phys. Rev. Lett.* **77**, 2778–2781 (1996).
4. Koshelev, A. E. & Vinokur, V. M. Dynamic melting of a moving lattice. *Phys. Rev. Lett.* **73**, 3580–3583 (1994).
5. Yaron, U. *et al.* Neutron diffraction studies of flowing and pinned magnetic flux lattices in 2H-NbSe₂. *Phys. Rev. Lett.* **73**, 2748–2751 (1994).
6. Bhattacharya, S. & Higgins, M. J. Dynamics of a disordered flux line lattice. *Phys. Rev. Lett.* **70**, 2617–2620 (1993).
7. Yaron, U. *et al.* Structural evidence for a two step process in the depinning of the superconducting flux-line lattice. *Nature* **376**, 753–755 (1995).
8. Duarte, A. *et al.* Dynamically induced disorder in the vortex lattice in 2H-NbSe₂. *Phys. Rev. B* **53**, 11336–11339 (1996).
9. Pardo, F. *et al.* Topological defects in the flux line lattice and their relationship to the critical current of a type II superconductor. *Phys. Rev. Lett.* **78**, 4633–4636 (1996).
10. Chen, M., Ho, T. H., Hui, S. W. & Pindak, R. Electron-diffraction study of free-standing liquid-crystal films. *Phys. Rev. Lett.* **59**, 1112–1115 (1987).
11. Bruinsma, R. & Nelson, D. R. Bond orientational order in smectic crystals. *Phys. Rev. B* **23**, 402–410 (1981).
12. Dominguez, D., Grønbech-Jensen, N. & Bishop, A. R. First order melting of a moving vortex lattice: effects of disorder. *Phys. Rev. Lett.* **78**, 2644–2647 (1997).
13. Marchevsky, M., Aarts, J. & Kes, P. H. Observation of correlated vortex flow in NbSe₂ with magnetic decoration. *Phys. Rev. Lett.* **78**, 531–534 (1997).
14. Scheidl, S. & Vinokur, V. M. Driven dynamics of periodic elastic media with disorder. *Phys. Rev. E* **57**, 2574–2593 (1998).
15. Balents, L., Marchetti, M. C. & Radzihovsky, L. Nonequilibrium steady states of driven periodic media. *Phys. Rev. B* **57**, 7705–7739 (1998).
16. Le Doussal, P. & Giarmarchi, T. Moving glass theory of driven lattices with disorder. *Phys. Rev. B* **57**, 11356–11403 (1998).

Correspondence and requests for materials should be addressed to D.J.B. (e-mail: djb@lucent.com).

A multiwavelength semiconductor laser

Alessandro Tredicucci, Claire Gmachl, Federico Capasso, Deborah L. Sivco, Albert L. Hutchinson & Alfred Y. Cho

Bell Laboratories, Lucent Technologies, 600 Mountain Avenue, Murray Hill, New Jersey 07974, USA

Many systems, such as atoms and molecules in gas mixtures, dye solutions and some solid-state materials, can exhibit simultaneous laser action at several wavelengths as a result of the excitation of several optical transitions¹. But semiconductor lasers are usually monochromatic because the electronic levels

are distributed in continuous energy bands². In order to achieve simultaneous lasing at several well-separated wavelengths, researchers have proposed³ combining different semiconductors with distinct bandgap energies in the active material. However, the difficulty of pumping different regions and of absorption of the shorter-wavelength light could be resolved only by using separated multiple resonators or by multisection injection devices^{4–7}. Here we report the realization of a single artificial semiconductor material with distinct optical transitions, which permits simultaneous multiwavelength laser action at mid-infrared wavelengths (6.6, 7.3 and 7.9 μm). This is achieved by tailoring the electronic states and electron relaxation times in the material, which is a superlattice layered structure. The laser has potential applications in sensors for trace-gas analysis.

Our structure is based on intersubband transitions, which connect quantized electronic states of the same conduction band. They resemble atomic transitions, because their joint density of states is fundamentally narrow (in principle delta-like)⁸, a unique feature which allows the design of materials transparent on both energy sides of an optical resonance. Intersubband injection lasers have

been realized some years ago with the application of the quantum cascade concept to electron transport⁹. The broad spectrum of accessible mid-infrared wavelengths (3.4–13 μm ; refs 10, 11), the room-temperature operation combined with their good reliability (ref. 12 and references therein), and the wide single-mode tuning range attainable with distributed-feedback resonators¹³ have made quantum-cascade lasers particularly suitable for gas sensing applications^{14,15}. The availability of a dual-wavelength laser would greatly simplify and improve their use in differential techniques relying on laser pulses of two different wavelengths (one resonant with an absorption line of the target gas and the other off-resonant); an example of such a technique is differential absorption lidar (light detection and ranging), one of the most sensitive spectroscopic methods of pollution monitoring¹⁶.

A tunable quantum-cascade laser exhibiting dual-wavelength operation at threshold was recently demonstrated¹⁷. However, the two optical transitions originated from separate sections of the same material biased at different voltages, to obtain two wavelengths via the Stark effect. Here we consider an active material intermediate between a superlattice structure with continuous energy minibands¹⁸, and multiple quantum wells with individual confined sub-bands¹⁹. In this material, the electronic states are delocalized over many wells and are grouped together in miniband-like manifolds, but their energy separation is sufficiently large when compared to the linewidth broadening that they can still appear as distinct features in the optical response. In a reciprocal-space picture, one can then speak of minibands with an energy dispersion $E(k_z)$, but the values of the wavevector k_z are quantized according to the number of periods in the superlattice. We have chosen eight-period superlattices, which results in discretized minibands composed of eight states. Eight well-defined optically allowed transitions, vertical in k -space and with relatively large oscillator strengths, are obtained between the first two minibands, each with a distinct frequency, and each corresponding to a different point of the Brillouin minizone (Fig. 1a). As discussed below, our structure is designed so that laser action can take place simultaneously on the transitions indicated as 4–1 and 3–2 in Fig. 1a. The energy separation between the upper states 4 and 3 is a critical parameter. It has to be large enough to make the difference between the emitted photon energies E_{4-1} and E_{3-2} considerably greater than

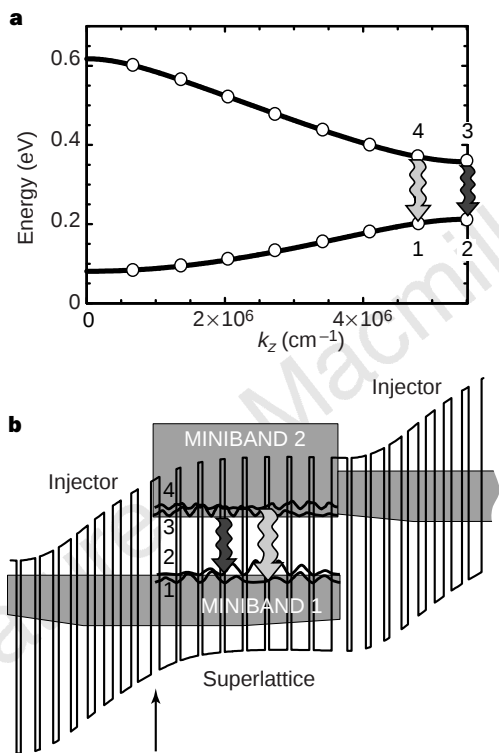


Figure 1 Conduction energy band diagrams. **a**, Dispersion of the first two minibands of a superlattice with 4.6-nm $\text{In}_{0.53}\text{Ga}_{0.47}\text{As}$ wells and 1.1-nm $\text{Al}_{0.48}\text{In}_{0.52}\text{As}$ barriers as calculated with a Kronig-Penney model in the reduced Brillouin zone scheme. The wavy arrows indicate the two points where laser action takes place. The circles indicate the eight individual states resulting from the finite number of superlattice periods. **b**, Conduction band profile of one superlattice region with the adjacent injectors at the design electric field of 29 kV cm^{-1} , as computed self-consistently using an iterative Schrödinger-Poisson calculation. The shaded regions indicate the energy and spatial extension of the manifold of miniband-like states. The moduli squared of the wavefunctions of the four relevant levels and the laser transitions are also shown in the figure. Starting from the barrier indicated by a vertical arrow the thicknesses in nanometres are (left to right): **2.2/4.3/1.1/4.5/1.1/4.6/1.1/4.6/1.1/4.6/1.1/4.6/1.1/4.6/1.1/4.6/4.0/1.1/3.7/1.2/3.4/1.4/3.1/1.6/2.7/1.9/2.3/2.3/2.0/2.7/1.7/3.3**, where the AlInAs barrier layers are in bold, and the underlined thicknesses represent the Si doped layers ($N_d = 5 \times 10^{17} \text{ cm}^{-3}$). We note that two of the superlattice wells have been narrowed from 4.6 nm to 4.3 nm and 4.5 nm, respectively, to compensate partially for the band profile at the edge of the superlattice not being completely flat.

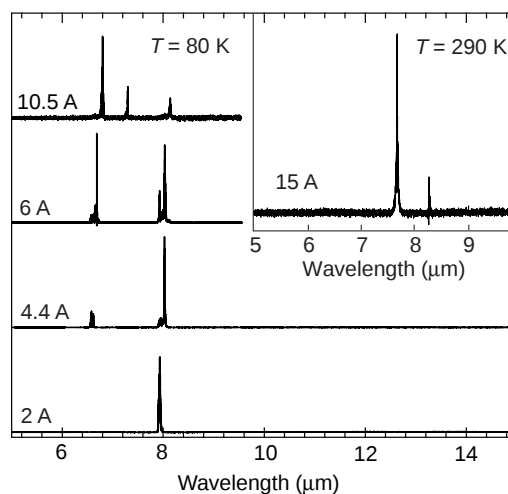


Figure 2 Laser emission spectra. Shown are emission spectra at 80 K of a laser stripe, 3 mm long and $17 \mu\text{m}$ wide, driven with current pulses of various amplitudes (width 50 ns, repetition rate 84.2 kHz for the 2 A, 4.4 A and 6 A spectra, 5 kHz for the 10.5 A and 15 A spectra), as detected in rapid-scan mode with a Fourier-transform infrared spectrometer. Inset, a room-temperature spectrum. The 50-ns pulse duration was chosen to avoid heating effects on the active material, but can be considered quasi-continuous wave from the point of view of carrier dynamics.

the broadening of the two transitions. On the other hand, because a sufficient population inversion has to be realized between 4 and 1, the relaxation rate of electrons from 4 to 3 must be kept small. This can be done by hindering the emission of optical phonons, which is the dominant scattering process; one approach is to design E_{4-3} to be less than the optical phonon energy E_{LO} . In our material system (InGaAs/AlInAs grown lattice-matched to InP) E_{LO} is ~ 34 meV, which poses stringent requirements on the maximum acceptable broadening. To minimize this broadening, we have used intrinsic superlattices (that is, undoped with a background doping concentration $N_d \approx 10^{15} \text{ cm}^{-3}$), which are kept electric-field-free under the applied bias by the use of special injection regions with an appropriate doping profile²⁰. This leads to luminescence linewidths as small as ~ 15 meV. The actual conduction-band profile of one superlattice with the two adjacent injectors is shown in Fig. 1b, together with the wavefunctions of the relevant levels. The miniband of each injector unit is realized so as to provide fast extraction of the electrons from the first miniband of the preceding superlattice active region and efficient injection into both states, 4 and 3, of the second miniband of the next superlattice. We calculate for the two transitions the wavelengths $\lambda_{41} = 6.4 \mu\text{m}$ and $\lambda_{32} = 7.9 \mu\text{m}$ and the transition matrix elements $z_{41} = 1.9 \text{ nm}$ and $z_{32} = 3.2 \text{ nm}$ (ref. 8). Relaxation through optical phonon emission into the other levels of the first miniband controls the lifetimes of states 1 and 2 (~ 0.2 – 0.3 ps), which are much shorter than the optical phonon scattering times $\tau_{32} \approx \tau_{41} = 10.5$ ps from 3 to 2 and from 4 to 1, thus ensuring population inversion (ref. 18; for the calculation of scattering times, see ref. 21). The total lifetimes of the upper states, neglecting the 4–3 relaxation processes, are computed to be $\tau_3 = 0.9$ ps and $\tau_4 = 1.1$ ps, as obtained by summing the scattering rates to all the states of the first miniband.

Twenty-five injector/superlattice stages were grown by molecular beam epitaxy on an n-type InP substrate. They were embedded in an n-doped InGaAs/AlInAs multilayer structure identical to the one used in ref. 20 to provide optical confinement and ohmic contacts. The wafers were then processed into mesa-etched ridge waveguide devices and cleaved into laser bars with the facets left uncoated. In Fig. 2 we show the emission spectra of one such laser at a temperature of 80 K. For current pulse amplitudes of 2 A only one laser line is visible at $\sim 7.9 \mu\text{m}$, but above 4 A a second line appears at $\sim 6.6 \mu\text{m}$, in good agreement with the expected wavelengths. Above 10 A, a third emitted wavelength appears at $\sim 7.3 \mu\text{m}$. The latter originates from a transition diagonal in k -space, which

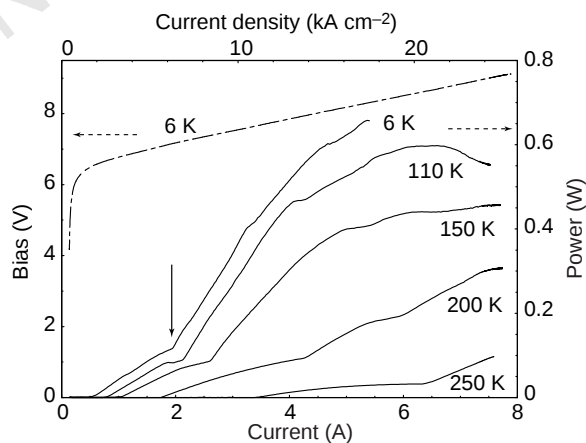


Figure 3 Light-current and voltage-current characteristics. Shown is peak output power as a function of peak drive current (50-ns pulses, 5-kHz repetition rate), measured with 50% collection efficiency from a single facet of a laser 2.26 mm long and $13.7 \mu\text{m}$ wide, using a room-temperature fast HgCdTe detector and a boxcar averager with an adjustable gate. The voltage-current (V - I) characteristic at 6 K is also shown.

acquires a large enough dipole matrix element only at high electric fields (see below). Above 250 K, lasing is achieved only at the two longer wavelengths (see the room-temperature spectrum in Fig. 2 inset). The emission wavelengths red-shift with temperature, and each line is actually composed of several longitudinal modes, as is typical for Fabry–Perot lasers. For spectroscopic purposes, however, these lines can be made single-mode by means of a double-period distributed-feedback resonator¹³. The light-current characteristics are plotted for various temperatures in Fig. 3. The laser shows the high peak optical powers (~ 1 W for the best device at 10 K) typical of superlattice quantum-cascade lasers¹⁸, but a pronounced kink appears in all curves (indicated by the solid arrow in the figure), marking the threshold for the onset of the second wavelength. A low temperature current–voltage curve is also shown in Fig. 3; laser emission starts at an applied bias of 6.5 V, in good agreement with the design electric field of 29 kV cm^{-1} in the active region. The threshold current densities (J_{th}) have been measured precisely for each single wavelength with the use of a monochromator, and, for the best device, we have obtained $J_{th}^{32} = 1.7 \text{ kA cm}^{-2}$ for the $7.9\text{-}\mu\text{m}$ transition (3–2) and $J_{th}^{41} = 5.8 \text{ kA cm}^{-2}$ for the $6.6\text{-}\mu\text{m}$ transition (4–1) at ~ 10 K. The usual expression of J_{th} for quantum-cascade lasers can be written for the discrete 3–2 transition as²²;

$$J_{th}^{32} = \frac{1}{\tau_3(1 - \tau_2/\tau_{32})} \left(\frac{\epsilon_0 \lambda_{32} n L_p \Delta_{32}}{4\pi q \Gamma z_{32}^2} (\alpha_m + \alpha_w) \right) \quad (1)$$

where ϵ_0 is the vacuum permittivity, $n = 3.27$ is the effective refractive index of the waveguide, Γ is the optical confinement factor, $L_p = 84.7 \text{ nm}$ is the length of one stage, $\Delta_{32} \approx 15$ meV is the measured full-width at half-maximum of the luminescence peak at cryogenic temperatures (see Fig. 4), $\alpha_m = 4.5 \text{ cm}^{-1}$ is the mirror loss, and α_w is the waveguide loss, which we have determined through the Hakki–Paoli technique²³ to be $\alpha_w = 21.5 \text{ cm}^{-1}$ at $7.9 \mu\text{m}$. These values would give a computed threshold of $\sim 1 \text{ kA cm}^{-2}$. However, we have to take into account the fact that the electron population is not concentrated only in level 3, as a considerable amount is also stored in level 4. By comparing the calculated with the experimentally measured J_{th}^{32} , we infer that only

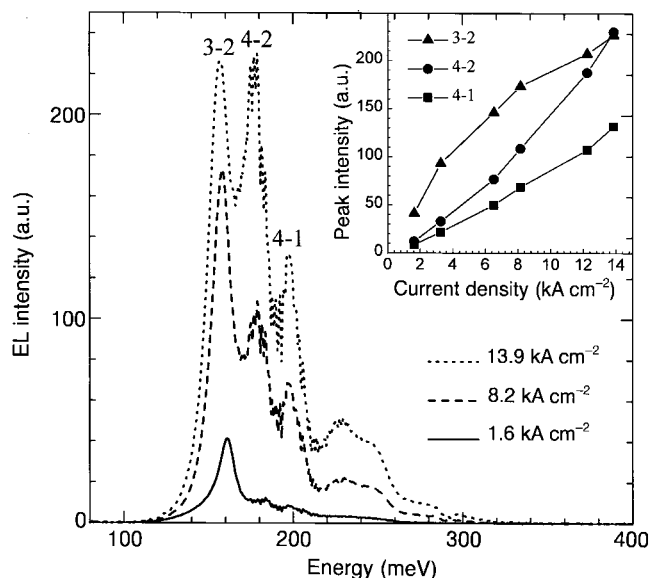


Figure 4 Electroluminescence (EL) spectra. Shown are pulsed EL spectra as measured from a round mesa ($125 \mu\text{m}$ diameter) with a Fourier-transform interferometer using a step-scan lock-in technique at a temperature of 5 K (a.u., arbitrary units). The fine structure superimposed on the main peaks above 170 meV is due to the absorption of water in the optical path. In the inset, the intensity of each peak is plotted as a function of current density.

60% of the electrons are actually injected into state 3. If we assume that the remaining 40% is then in level 4, and that $\alpha_w \propto \lambda^2$, a dependence characteristic of free-carrier absorption, we obtain a theoretical estimate for J_{th}^{41} of 3.8 kA cm⁻². The discrepancy with the experimental value probably originates from electrons in higher states of the upper miniband (see also the data of Fig. 4) and from direct scattering from the injector to the lower miniband, which leads to non-unity injection efficiency into 3 and 4. At high bias, the doping charges in the injector are no longer sufficient to keep the superlattices free of electric field. Therefore, the superlattice translation and reflection symmetries are spoiled and transitions which are suppressed in an ideal superlattice can acquire large dipole matrix elements. In particular, 4–2 (which is diagonal in *k*-space) is the one responsible for the $\lambda = 7.3 \mu\text{m}$ laser emission.

More information on the actual behaviour of our structure can be obtained by analysing the luminescence spectra (Fig. 4). Three peaks with a width of ~ 15 meV are clearly identified, in correspondence with the energy of each of the lasing transitions. Although the relative intensity of the peaks depends on the wavelength response of the experimental set-up, it is evident from Fig. 4 inset that the intensity of the 4–2 peak increases with current density more rapidly than that of the other peaks, which we expect in view of its oscillator strength increasing (and the other two decreasing) with bias. Emission at larger energies (≥ 220 meV) is also detected, due to transitions from states higher in the miniband.

We note that a complete localization of the states over one or two wells induced by the penetration of the electric field²⁴ would give rise to completely different spectra, and is not compatible with the smooth and monotonic current–voltage curves we observe. Furthermore, the density of carriers stored in the upper miniband at threshold is estimated from the lifetimes to be $\sim 6 \times 10^{10} \text{ cm}^{-2}$, far less than the sheet density of doping in the injectors, which ensures negligible space-charge electric fields caused by the injection. To estimate the threshold for lasing on the 4–2 transition, we cannot use equation (1), because the population n_4 in level 4 is, to a first approximation, locked to its value at the threshold for the 4–1 transition. The threshold for 4–2 lasing is instead reached by a different mechanism (laser action by oscillator strength tuning¹⁷). As the penetration of the electric field into the superlattice increases, the dipole matrix element z_{42} is enhanced until the gain equals the optical losses (laser threshold condition). Roughly speaking then, J_{th}^{42} should be determined by the condition that the applied voltage is such that $z_{42} \approx z_{41}$, which occurs at 7.3–7.4 V, in fair agreement with the measured 7.7 V. □

Received 17 July; accepted 7 October 1998.

1. Silvast, W. T. *Laser Fundamentals* (Cambridge Univ. Press, 1996).
2. Yariv, A. *Optical Electronics* (Saunders College Publishing, Philadelphia, 1991).
3. Ikeda, S., Shimizu, A. & Hara, T. Asymmetric dual quantum well laser-wavelength switching controlled by injection current. *Appl. Phys. Lett.* **55**, 1155–1157 (1989).
4. Beernik, K. J., Thornton, R. L. & Chung, H. F. Low threshold current dual wavelength planar buried heterostructure lasers with close spatial and large spectral separation. *Appl. Phys. Lett.* **64**, 1082–1084 (1994).
5. Garcia, J. Ch. *et al.* Epitaxially stacked lasers with Esaki junctions: a bipolar cascade laser. *Appl. Phys. Lett.* **71**, 3752–3754 (1997).
6. Dutta, N. K. *et al.* InGaAsP closely spaced dual wavelength laser. *Appl. Phys. Lett.* **48**, 1725–1726 (1986).
7. Pellandini, P. *et al.* Dual-wavelength laser emission from a coupled semiconductor microcavity. *Appl. Phys. Lett.* **71**, 864–866 (1997).
8. Sirtori, C., Capasso, F., Faist, J. & Scandolo, S. Nonparabolicity and a sum rule associated with bound-to-bound and bound-to-continuum intersubband transitions in quantum wells. *Phys. Rev. B* **50**, 8663–8674 (1994).
9. Faist, J. *et al.* Quantum cascade laser. *Science* **264**, 553–556 (1994).
10. Faist, J. *et al.* Short wavelength ($\lambda \sim 3.4 \mu\text{m}$) quantum cascade laser based on strained compensated InGaAs/AlInAs. *Appl. Phys. Lett.* **72**, 680–682 (1998).
11. Gmachl, C. *et al.* Long wavelength ($\lambda \sim 13 \mu\text{m}$) quantum cascade lasers. *Electron. Lett.* **34**, 1103–1104 (1998).
12. Capasso, F., Faist, J., Sirtori, C. & Cho, A. Y. Infrared (4–11 μm) quantum cascade lasers. *Solid State Commun.* **102**, 231–236 (1997).
13. Gmachl, C. *et al.* Continuous-wave and high-power pulsed operation of index-coupled distributed feedback quantum cascade lasers at $\lambda \approx 8.5 \mu\text{m}$. *Appl. Phys. Lett.* **72**, 1430–1432 (1998).
14. Namjou, K. *et al.* Sensitive absorption spectroscopy with a room-temperature distributed-feedback quantum-cascade laser. *Opt. Lett.* **23**, 219–221 (1998).
15. Sharpe, S. W. *et al.* High-resolution (Doppler-limited) spectroscopy using quantum-cascade distributed-feedback lasers. *Opt. Lett.* **23**, 1396–1398 (1998).

16. Sigrist, M. W. *Air Monitoring by Spectroscopic Techniques* (Wiley, New York, 1994).
17. Faist, J. *et al.* Laser action by tuning the oscillator strength. *Nature* **387**, 777–782 (1997).
18. Scamarcio, G. *et al.* High-power infrared (8-micrometer wavelength) superlattice lasers. *Science* **276**, 773–776 (1997).
19. Dingle, R., Gossard, A. C. & Wiegmann, W. Direct observation of superlattice formation in a semiconductor heterostructure. *Phys. Rev. Lett.* **34**, 1327–1330 (1975).
20. Tredicucci, A. *et al.* High-power inter-miniband lasing in intrinsic superlattices. *Appl. Phys. Lett.* **72**, 2388–2390 (1998).
21. Ferreira, R. & Bastard, G. Evaluation of some scattering times for electrons in unbiased and biased single- and multiple-quantum-well structures. *Phys. Rev. B* **40**, 1074–1086 (1989).
22. Sirtori, C. *et al.* Continuous wave operation of midinfrared (7.4–8.6 μm) quantum cascade lasers up to 110 K temperature. *Appl. Phys. Lett.* **68**, 1745–1747 (1996).
23. Hakki, B. W. & Paoli, T. L. Gain spectra in GaAs double-heterostructure injection lasers. *J. Appl. Phys.* **46**, 1299–1306 (1975).
24. Mendez, E. E., Agullo-Rueda, F. & Hong, J. M. Stark localization in GaAs–GaAlAs superlattices under an electric field. *Phys. Rev. Lett.* **60**, 2426–2429 (1988).

Acknowledgements. This work was partially supported by the Office of Naval Research and by DARPA/US Army Research Office.

Correspondence and requests for materials should be addressed to E.C. (e-mail: fc@lucent.com).

Delaminated zeolite precursors as selective acidic catalysts

A. Corma*, V. Fornes*, S. B. Pergher*, Th. L. M. Maesen† & J. G. Buglass‡

* Instituto de Tecnología Química (U.P.V.–C.S.I.C.), Universidad Politécnica de Valencia, Avenida de los Naranjos, s/n, 46022 Valencia, Spain

† Zeolyst International, PQ Research and Development Center, Conshohocken, Pennsylvania 19428–2240, USA

‡ Shell International Chemicals, B. V., Shell Research and Technology Center, 1030 BN Amsterdam, The Netherlands

Heterogeneous catalysis is important in fine-chemical and pharmaceutical manufacture and in petroleum refining¹. Many of the catalysts used by these industries are based on aluminosilicates, which combine high stability with excellent activity in acid-mediated reactions². Within this class of material, zeolites—microporous crystalline aluminosilicates with three-dimensional framework structures—have attracted particular attention: they are significantly more active than the layered structures (clays)³ and mesoporous structures (whose walls are amorphous)^{4,5}, and they impart shape selectivity on the reaction products. The selectivity arises from the fact that the catalytically active acid sites have to be accessed through uniformly sized pores and voids, imposing size constraints on the accessibility to reactants and the nature of the intermediates and products¹. Providing access for larger molecules to the catalytic sites would expand the range of reactions that zeolites can catalyse. But attempts to increase the pore size of zeolites⁶ have met with only limited success. Here we describe an approach in which a layered zeolite precursor^{7,8} is delaminated, in much the same way as the layered structure of a clay may be unbound. The result is an aluminosilicate whose zeolite-type catalytic sites are contained within thin, readily accessible sheets. Performance tests show that the delamination process improves the accessibility of the catalytic sites without affecting their activity. We expect that our approach could be adapted to other layered zeolite precursors^{5,7–11}, thus paving the way to exfoliated zeolite precursors as a new class of catalysts.

Attempts to enlarge the access routes to the catalytic sites of zeolites have so far focused on increasing pore sizes, with the largest pore size achieved to date being 0.7 nm (ref. 12). Our approach, in contrast, exploits the layered nature of one of an increasing number of zeolite precursors that have been reported^{8–11}. By delaminating MCM-22(P), the precursor of both MCM-22 and ERB-1 zeolites (which share the MWW structure), we have succeeded in synthesizing a more readily accessible aluminosilicate, designated ITQ-2

(Instituto de Tecnología Química)¹³. The inherent structural complexity of ITQ-2 poses many problems when attempting to provide an all-encompassing, quantitative model of its structure. However, based on the structure of MCM-22 and the proposed structure for MCM-22(P)^{7,14}, it seems likely that ITQ-2 consists of thin sheets (~2.5 nm thick) with an extremely high, and structurally well defined, external surface area ($\geq 700 \text{ m}^2 \text{ g}^{-1}$). These sheets should consist of an hexagonal array of 'cups' that penetrate into the sheet from both sides. These cups would have an aperture of ~0.7 nm, formed by a 12-membered ring (a ring of $12 \equiv \text{Si}-\text{O}-\text{Si} \equiv$ units). We estimate that these cups are ~0.7 nm deep. The cups meet at the centre of the layer, forming a double 6-ring window that connects the cups, bottom to bottom. As a result, a smooth, 10-membered ring channel system runs in between the cups, inside the sheet.

ITQ-2 was made by first synthesizing MCM-22(P)⁷, a zeolite precursor for the MWW-type structure⁷, which consists of inorganic layers connected together by a layer of organic material (hexamethyleneimine, HMI). Typically 0.23 g of sodium aluminate (56% Al_2O_3 , 37% Na_2O , Carlo Erba, Milan) and 0.81 g of sodium hydroxide (98%, Prolabo, Barcelona) were dissolved in 103.45 g of distilled water, after which 6.35 g HMI and 7.86 g of silica Aerosil 200, Degussa, Frankfurt) were added consecutively. The mixture was stirred vigorously for 30 minutes at room temperature, following by 11 days in a stirred PTFE-lined stainless-steel autoclave at 408 K under autogenous pressure. The crystalline product was filtered, washed with distilled water until $\text{pH} < 9$, and the filter cake mixed with water to give a slurry with 20 wt% solids.

The next step involved swelling of the solids, and was carried out

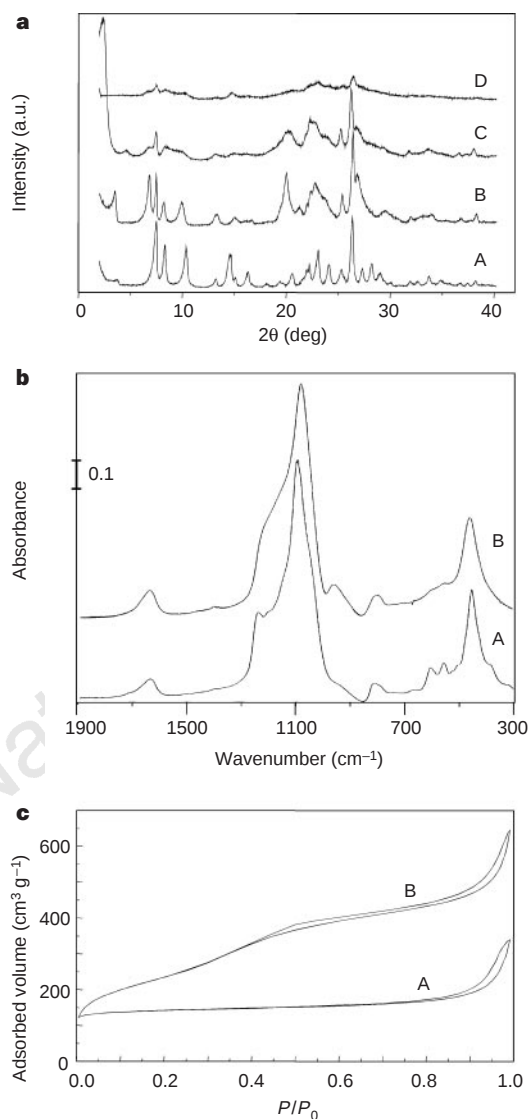


Figure 1 Physical characterization of the MWW-type zeolite and ITQ-2. **a**, X-ray diffraction patterns (Cu $K\alpha$ radiation, $\lambda = 1.540598 \text{ \AA}$) of the MWW-type zeolite (trace A), MCM-22(P), which contains well-ordered alternate layers of ITQ-2 and organic HMI (trace B), swollen MCM-22(P) (trace C) and delaminated ITQ-2 layers (trace D). The diffraction patterns are offset for clarity. **b**, Infrared spectra of the framework region of the MWW-type zeolite (spectrum A) and ITQ-2 (spectrum B). The spectra are offset for clarity. **c**, Nitrogen adsorption isotherms of the MWW-type zeolite (isotherm A) and ITQ-2 (isotherm B). P/P_0 is the partial pressure as a fraction of the saturation pressure (P_0).

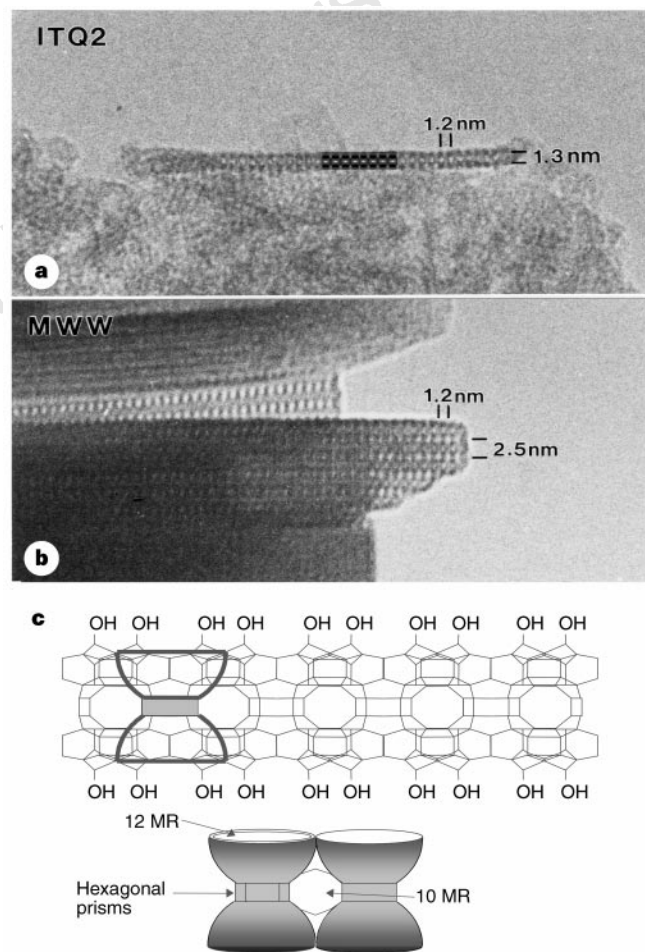


Figure 2 Structural characterization of ITQ-2 and the MWW-type zeolite. **a**, Transmission electron microscopy (TEM) image of an ~2.5-nm-thick ITQ-2 layer, viewed along the 10-ring channels, perpendicular to the c axis. The small, dark insert (centre) is a simulation at a defocus of -60 nm. **b**, TEM image of the MWW-type zeolite sample, viewed perpendicular to the c direction. The fusion of the layers along the c direction to form supercages is clearly visible. **c**, Proposed structure for the ITQ-2 layer, as employed in the simulation, showing the characteristic 10-membered ring (10 MR) separating arrays of 'chalices' perpendicular to (001), with an artist's impression of one of the chalices included. The bottom illustration is an artist's impression of two fused chalices, each made of two 'cups', connected by a non-shared 6-membered ring at the bottom, and with a 12-membered ring (12 MR) at their open top. The two fused chalices enclose a 10-membered ring, which forms parts of the channel running between the cups inside the sheet.

by refluxing a mixture of 27 g of slurry with 105 g of an aqueous solution of 29 wt% hexadecyltrimethylammonium bromide and 33 g of an aqueous solution of 40 wt% tetrapropylammonium hydroxide for 16 hours at 353 K. The completion of the swelling can be monitored by X-ray diffraction, which shows an increase in the distance between the layers from 2.7 nm to ~ 4.5 nm (Fig. 1 trace C). The layers are forced apart by placing the slurry in an ultrasound bath (50W, 40 kHz) for one hour. A few drops of concentrated hydrochloric acid are then added until the pH is below 2, and the solids are collected by centrifuging. Organic material is then removed by calcination of the solids at 813 K, yielding ITQ-2. A portion of the MCM-22(P) starting material was calcined, and used as the MWW-type zeolite reference sample for the following characterization work.

Figure 1a shows the X-ray diffraction patterns of the different intermediates produced during the preparation of ITQ-2. ITQ-2 (Fig. 1 trace D) does not show the 001 and 002 peaks at $2\theta = 3-7^\circ$ which is consistent with the proposed structures; that is, ITQ-2 does not have the regular array of layers, with its characteristic 2.5-nm periodicity, typical of the MWW topology. Comparing the X-ray diffraction patterns of ITQ-2 with that of an MWW-type zeolite of similar silica-to-alumina ratio (Fig. 1 trace A) reveals that the high-angle peaks are much broader for ITQ-2, which is indicative of a reduction in the size of the domain of coherent scattering; that is, exfoliation of the MCM-22(P) has significantly reduced the long-range order in the new structure. Nonetheless, ITQ-2 and the MWW-type zeolite reference sample yield similar infrared spectra in the region representing the framework vibrations (Fig. 1b), implying that the short-range order in ITQ-2 has remained virtually unchanged.

ITQ-2 and the MWW-type zeolite both show a number of infrared-active bands in the $4,000-300\text{ cm}^{-1}$ range (Fig. 1b). Of particular interest is the band around 950 cm^{-1} , which corresponds to silanol groups. As shown in the figure, this band is more intense in ITQ-2, suggesting that the material contains a larger number of silanol groups than the zeolite. Infrared spectra of the OH stretching region, which allow the identification of silanols at $3,745\text{ cm}^{-1}$, corroborate this conclusion (data not shown). Similarly, the ^{29}Si magic-angle-spinning (MAS) NMR spectra of ITQ-2 and MWW-type zeolite (not shown) are comparable except for the peak at 119 p.p.m. corresponding to Si–O–Si bridges between the layers in the MWW-type zeolite. In the ^{29}Si MAS NMR spectrum of ITQ-2 this peak is no longer present and, instead, an intense peak appears at 102 p.p.m. corresponding to (terminal) $[\text{Si}-\text{O}]_3\text{-Si-OH}$ groups. The ratio of the intensity of the peak at 102 p.p.m. to that at 119 p.p.m. appears to be proportional to the degree of delamination as assessed by external surface area (see below). This supports the suggestion that the delamination has resulted in the ITQ-2 layers terminating in silanol groups, whereas in the MWW-type zeolite

these silanol groups are condensed into Si–O–Si bridges between the layers.

The effects of the delamination are also evident in the nitrogen absorption measurements. Figure 1c shows the N_2 adsorption isotherm of an MWW-type zeolite, compared to that of an ITQ-2 sample of the same silicon-to-aluminium ratio. The results for the ITQ-2 sample show a significantly higher mesopore volume and an external surface area of $\sim 700\text{ m}^2\text{ g}^{-1}$, which is about twice that of the MWW-type zeolite or of MCM-56, a related zeolite¹⁵. This is consistent with the ITQ-2 sheets being arranged randomly and mainly edge-to-face, as opposed to the face-to-face orientation prevalent in MCM-56 zeolite.

Transmission electron microscopy (TEM) is used to visually inspect the ITQ-2 sheets. Wherever the ITQ-2 sheets curl up, TEM is able to distinguish between the regular arrays of exposed $0.7 \times 0.7\text{ nm}$ cups and the 10-ring channel system in between them (Fig. 2a). A TEM image of the MWW-type zeolite is shown in Fig. 2b for comparison, clearly showing the bonding of cups to form supercages. Based on the proposed ITQ-2 structure (Fig. 2c), simulations of the TEM image, included in Fig. 2a, were performed using the high-resolution simulation program Cerius2 (ref. 16). The results are consistent with the observed TEM image, thereby supporting the proposed ITQ-2 structure. The results of a more-detailed TEM study on a through-focus image series are available as Supplementary Information. Most of the ITQ-2 layers are bent and curled (similar to clays), interfering with a detailed X-ray verification of the proposed structure.

The assertion that this 10-membered ring channel system is preserved during the delamination process is corroborated by temperature programmed desorption–mass spectrometry (TPD-MS) measurements used to follow the decomposition of the HMI in the zeolite precursor samples. In the case of the MWW-type precursor, two desorption peaks are observed (at 578 K and 708 K). These can be attributed to decomposing HMI escaping through the inter-layer channels and the 10-membered ring (intra-layer) channels, respectively. In contrast, TPD-MS on the ITQ-2 precursor yields only one desorption peak (at 708 K)—attributable to the decomposing HMI escaping through the intra-layer 10-membered ring channels.

The catalytic potential of ITQ-2 was illustrated by means of a fixed-bed, small-scale catalytic-cracking test. Using *n*-decane as model feed, 0.5 g of catalyst is diluted with 2.5 g of silica. The time on stream is 60 seconds (and contact times are between 8.4 and 25.2 g catalyst per g feed per s) at 770 K, followed by stripping with nitrogen for 15 minutes. The conversions of the reactant observed at different contact times are then used to calculate first-order rate constants (Fig. 3). The similar rate constants for *n*-decane cracking indicate that the MWW-type zeolite sample and ITQ-2 (both with a silicon-to-aluminium ratio of 50) have similar activities; that is, the delamination process does not seem to have destroyed the zeolite character of the catalytic sites in ITQ-2. We note, however, that the selectivity (ratio of desired to undesired products) of ITQ-2 and the MWW-type zeolite are remarkably different in this test: ITQ-2 catalyses the formation of significantly more liquid and less gaseous product than does the MWW-type zeolite. (At a conversion level of 26%, ITQ-2 catalyses the conversion of 9.5% and 15% of *n*-decane into liquids and gas, respectively. For the MWW-type zeolite, in contrast, the values are 7% and 17%.) The relative amount of gas formation provides a measure of the extent to which the initial liquid products have been further cracked in consecutive reactions, and is a function of catalytic-site accessibility. The reduced gas formation observed with ITQ-2 suggests that fewer consecutive reactions have taken place; that is, the initial products seem to be able to diffuse more rapidly out of ITQ-2 than out of the MWW-type sample.

The benefits of the greater site accessibility and smaller diffusion path in ITQ-2 can be fully appreciated when larger molecules are

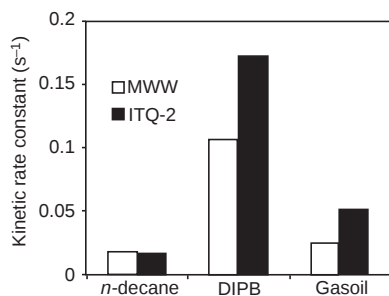


Figure 3 Comparison of the first-order kinetic rate constants for the cracking of *n*-decane, 1,3-diisopropylbenzene (DIPB) and vacuum gasoil over the MWW-type zeolite and ITQ-2. The reactants are cracked for periods of 60 s (see text), and the observed conversions then used to calculate first-order rate constants.

used as feedstocks, such as di-isopropylbenzene and vacuum gasoil. The ITQ-2 sample is more active (Fig. 3) as well as more selective than the MWW-type zeolite, yielding more of the valuable gasoline and diesel products and less gas and coke. (At 48% conversion, gasoline and diesel constitute two-thirds of the products obtained with ITQ-2, but less than half of the products obtained with the MWW-type zeolite.) Whereas the selectivity advantage can be attributed to the fact that the desorption of product molecules, before being destroyed by consecutive reactions, is easier in ITQ-2, the activity advantage is due to the increased access to the catalytic sites for larger molecules. □

Received 17 October 1997; accepted 1 October 1998.

- Maxwell, I. E. & Stork, W. H. J. in *Introduction to Zeolite Science and Practice* (eds. van Bekkum, H. et al.) 571–630 (Elsevier, Amsterdam, 1991).
- Corma, A. Inorganic solids and their use in acid-catalyzed hydrocarbon reactions. *Chem. Rev.* **95**, 559–614 (1995).
- Burch, R. Pillared clays. *Catal. Today* **2**, 185–367 (1988).
- Corma, A., Grande, M. S., Gonzalez-Alfaro, V. & Orchilles, A. V. Cracking activity and hydrothermal stability of MCM-41 and its comparison with amorphous silica-alumina and USY zeolite. *J. Catal.* **159**, 357–382 (1996).
- Corma, A., Fornés, V., Navarro, M. T. & Pérez-Pariente, J. Acidity and stability of MCM-41 crystalline aluminosilicates. *J. Catal.* **148**, 569–575 (1994).
- Barrer, R. M. *Hydrothermal Chemistry of Zeolites* (Academic, London, 1982).
- Roth, W. J. et al. in *Catalysis by Microporous Materials* (eds Beyer, H. K. et al.) 301–308 (Elsevier, Amsterdam, 1991).
- Lawton, S. et al. Zeolite MCM-49: a 3-dimensional MCM-22 analog synthesized by in-situ crystallization. *J. Phys. Chem.* **100**, 3788–3798 (1996).
- Whittam, Th. V. Zeolites. Eur. Patent 0054364 (1982).
- Schreyeck, L., Caullet, P., Mouguel, J. C., Guth, J. L. & Marier, B. Prefer: a new layered (aluminosilicate precursor of FER-type zeolite. *Microporous Mater.* **6**, 259–271 (1996).
- Lohse, U. et al. Synthesis of zeolite-beta. Formation of zeolite-beta and titanium-beta via an intermediate layer structure. *J. Chem. Soc. Faraday Trans.* **93**, 505–512 (1997).
- Freyhardt, C. C., Tsapatsis, M., Lobo, R. F., Balkus, K. J. Jr & Davis, M. E. A high-silica zeolite with a 14-tetrahedral-atom pore opening. *Nature* **381**, 295–298 (1996).
- Corma, A., Fornés, V. & Pergher, S. B. EP (Spanish patent) 9502188 (1996); PCT-WO Patent 97/17290 (1997).
- Leonowicz, M. E., Lawton, J. A., Lawton, S. L. & Rubin, M. K. MCM-22: a molecular sieve with two independent multidimensional channel systems. *Science* **264**, 1910–1913 (1994).
- Fung, A. S., Ford, C. H., Lawton, S. L. & Roth, W. J. Synthetic layered MCM-56, its synthesis and uses. US Patent 5362697 (1994).
- Cerius2 (Biosym/Molecular Simulations Inc., San Diego, 1995).

Supplementary information is available on Nature's World-Wide Web site (<http://www.nature.com>) or as paper copy from the London editorial office of Nature.

Acknowledgements. We thank C. H. Roemkens for comments on the manuscript, and R. Haswell for assistance with the supplementary TEM work.

Correspondence and requests for materials should be addressed to A.C. (e-mail: acorma@itq.upv.es).

Spontaneous movement of ions through calcite at standard temperature and pressure

S. L. S. Stipp*, J. Konnerup-Madsen*, K. Franzreb†‡, A. Kulik§ & H. J. Mathieu†

* Geologisk Institut, Københavns Universitet, Øster Voldgade 10, DK-1350 København K, Denmark

† Laboratoire Metallurgie Chimique, Matériaux, § Institut Genie Atomique, Physique, Ecole Polytechnique Fédérale de Lausanne, CH-1015 Lausanne, Switzerland

At the resolution limits of traditional geochemical techniques, there is little evidence to challenge the common assumptions that, under the Earth's ambient surface conditions, dry calcite is static and that the bulk mineral behaves as a closed system. Solid-state diffusion has been recognized at elevated temperatures^{1–3}, but it has always been assumed that diffusion in carbonate minerals is negligible under standard conditions^{4,5}. There is, however, some

evidence to the contrary. More than 30 years ago, the ⁴⁵Ca diffusion coefficient was estimated to be $\sim 8 \times 10^{-20} \text{ cm}^2$ (ref. 6) and, more recently, we have demonstrated movement of adsorbed Cd²⁺ and Zn²⁺ into bulk calcite at rates of tens of nanometres over weeks to months (refs 7, 8). Here we present evidence that monovalent ions, Na⁺, K⁺ and Cl[−], originating from fluid inclusions, accumulate in crystallites on the surface of calcite. This process is spontaneous at the Earth's surface conditions, in air. The results show that calcite under standard conditions does not always behave as a closed system, which is a critical assumption in the use of isotope ratios, trace-element distribution and fluid-inclusion composition for interpretations of palaeoclimate, geochronology or petrogenesis. Moreover, calcite's uptake capacity for contaminants in environmental systems is probably higher than current models predict, because surface sites are constantly renewed by ionic mobility.

Although much is already understood about calcite structure⁹ and solubility¹⁰, little is known about its surface behaviour. Surface analytical techniques, which are common tools in physics and materials science, provide molecular-level observations from the topmost atomic layers of solids, so they have much to offer studies of geochemical processes that are too slow to yield macroscopic changes on a human timescale. When adsorbed cations were seen to move into calcite, an important question arose: was movement by solid-state diffusion or simply by surface diffusion into microfractures? Atomic force microscopy (AFM)¹¹, a surface-sensitive method, provides physical information about the topmost atomic layer, with horizontal resolution of about 0.1 nm and vertical resolution of 0.01 nm. AFM images of fresh calcite surfaces are shown in Fig. 1a–c; at this scale, micrometre- or nanometre-scale fractures would be visible if they were present. Examination of hundreds of samples proved that fresh surfaces always have atomically flat terraces separated by cleavage steps of various heights. Although millimetre-scale fractures are sometimes visible by eye or light microscope, neither micrometre- nor nanometre-scale fractures have ever been observed. Occasionally, edges of a fluid inclusion are visible if cleavage happens to intersect one, but we note that the smallest fluid inclusions visible with the usual light microscope are at least four orders of magnitude larger than the limits of AFM resolution.

Investigation of cleaved fragments during exposure to air for several hours has proved that the calcite surface is dynamic¹². The rate of change depends on relative humidity, but within minutes, material disappears from step edges, from calcite particles left after cleavage and from defect sites in terraces. New material accumulates at step edges, in monolayer-high bumps on terraces, or in dendritic overgrowths several atomic layers thick. Compare a freshly cleaved surface (Fig. 1c) to the same sample imaged a few hours later (Fig. 1d). From AFM tip response to surface forces, we know that a thin water film adsorbs from the humidity in air¹². Synchrotron X-ray reflectivity¹³ has indicated one such film to be ~ 20 water molecules thick (5–6 nm). It seems that calcite dissolves in this adsorbed film until saturation is reached and dynamic equilibrium induces the surface, which we see as dry, to recrystallize. This surface mixing process undoubtedly entrains adsorbed ions and releases foreign material trapped within the near-surface but we have not yet observed modification extending beyond about 10 atomic layers (~ 3 nm), which is much less than the depth needed to explain the decrease in surface concentration of Cd²⁺ and Zn²⁺.

On samples of calcite exposed to air for many months, surface features formed that ranged in height and width from a few nanometres to $\sim 1 \mu\text{m}$ (Fig. 1e). AFM is excellent for observing physical features on a surface but it cannot provide chemical information. A complementary technique, time-of-flight secondary ion mass spectroscopy, TOF-SIMS¹⁴, can give chemical maps of the very top atomic layers. Figure 2 presents TOF-SIMS images of the same surface shown in Fig. 1e. Using natural and etched markers on various samples, we have been able to prove¹⁵ that the spots with

‡ Present address: Surface Science Western, University of Western Ontario, London, Ontario N6A 5B7 Canada.

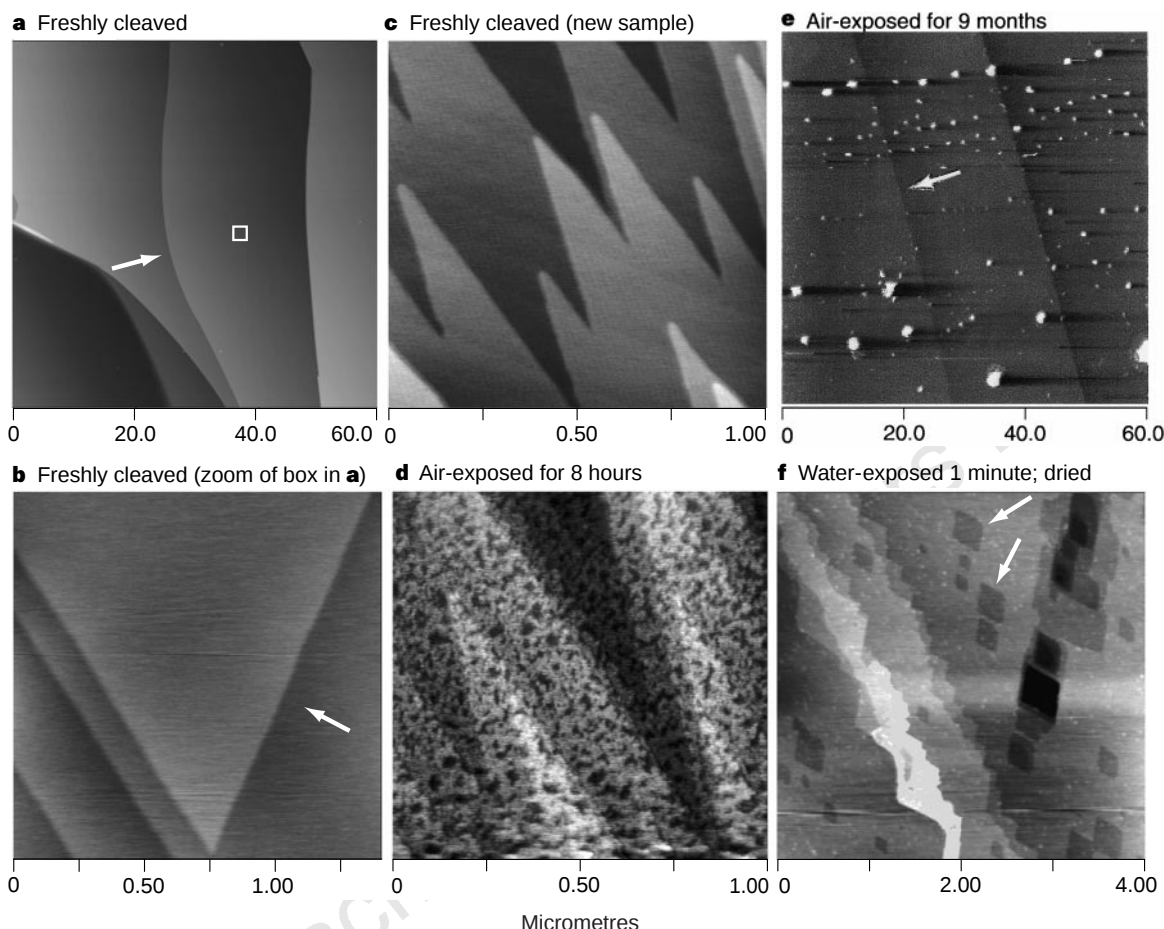


Figure 1 Atomic force microscopy images, taken in air, of calcite surfaces. We note that lateral and vertical scales vary between images. AFM uses a piezoelectric element to scan a sample in sub-nanometre increments under a sharp tip mounted on a cantilever with a very light spring constant¹¹. When the tip comes close to a surface, atomic-level forces (nanonewton range) cause cantilever deflections, which are recorded as a function of location on the surface, giving an image resembling surface topography. Topographically higher areas appear lighter, and darker areas lower. Dark and bright shadows on **e** and **f** are artefacts. **a**, Calcite, immediately following cleavage; step height at arrow, ~40 nm. **b**, Higher-resolution image taken at the square in **a**; step height at the arrow is 0.3 nm, one calcite atomic layer. Freshly cleaved surfaces are atomically flat, often over many micrometres; no holes or micro-fractures can be found. **c**, A new sample, immediately following cleavage (step height, 0.3 nm), and **d**, 8 hours later. Calcite surfaces rearrange spontaneously in the liquid film that

adsorbs from the humidity in air. Holes are one atomic layer (0.3-nm) deep. Examination of many samples shows the influence depth of this rearranging process is always less than 10 atomic layers. **e**, After exposure to air for 9 months, crystallites have formed, with width and height ranging from a couple of micrometres to a few nanometres. Step height at arrow, ~90 nm. This image compares with **a** in both horizontal and vertical scale. **f**, Another fresh sample which was exposed to deionized (Milli-Q) water for one minute and then dried. Rhombohedral etch pits (arrows), which are one atomic layer deep, are typical for dissolution of well-crystallized calcite. Each pit spreads across the surface, as thermodynamics predicts, before a new, one-layer-deep pit nucleates and spreads again. In contrast, the dark holes aligned along one cleavage direction are much deeper. They result from enhanced dissolution along a set of linear defects projecting into the crystal face.

higher relief are also enriched in monovalent ions, while the flatter surface maintained its original CaCO_3 composition. Now the question became: do the crystallites result from airborne contamination, or from material originating within the crystal?

All experiments used single-crystals of Iceland spar from Chihuahua, Mexico (atomic absorption spectroscopy, AAS, and inductively coupled plasma atomic emission spectroscopy, ICP-AES analyses: >99% CaCO_3). They were not thin-sectioned; fresh surfaces were prepared by gently scoring with a scalpel along one cleavage direction. The requirements for ultra-high vacuum, necessary for TOF-SIMS, preclude the use of epoxy, resin, glue and so on, so samples were always held mechanically, eliminating the possibility of contamination from mounting media. All tools, gloves and surfaces were dust- and grease-free, and reproducibility was proven using especially stringent precautions. Surfaces imaged with TOF-SIMS immediately following cleavage are consistently homogeneous, confirming that only elements expected for calcite are present and that monovalent ions are initially absent (Fig. 3a–c).

After ageing for several months, crystallites, aligned along one cleavage direction, are observed all over some of the samples (Fig. 2), in broad bands on others (Fig. 3d–f), and not at all on others. All samples were prepared and stored in the same way, in Lausanne, Switzerland, far from airborne sea-salt, in boxes closed to dust but open to atmospheric pressure. In bands lacking crystallites, measured surface concentrations of Na^+ , K^+ and Cl^- are at background levels, evidence against air-derived contamination.

Both AFM and TOF-SIMS reveal crystallites aligned along one of the intersecting cleavage directions and elongated along the other. AFM images with progressively higher resolution show progressively smaller crystallites to a limit of a few nanometres. Crystallite shape and distribution are reminiscent of fluid inclusions, which can range in size from visible to only a few molecules¹⁶. Optical microscope examination demonstrates that indeed, some samples have fluid inclusions and that they are aligned along one cleavage direction and elongated along another (Fig. 4); often, they occur in bands adjacent to zones completely free of them. Samples with visible fluid inclusions

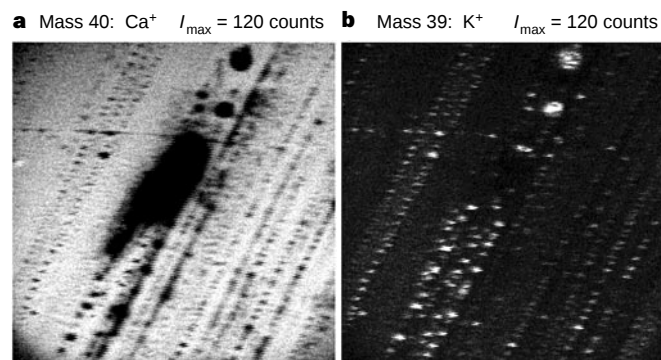


Figure 2 Positive time-of-flight secondary ion mass spectrometry maps for Ca and K from the same aged surface shown in Fig. 1e. TOF-SIMS uses a pulsed ion beam (in this case, $^{69}\text{Ga}^+$) to eject pulses of secondary ions, whose masses are recorded as a function of time in a mass analyser¹⁴. Mass resolution is better than 0.01 atomic mass units and spacial resolution is $\leq 1 \mu\text{m}$. Image width is $180 \mu\text{m}$. " I_{max} " represents the maximum secondary ion intensity for the brightest areas: dark areas have lower count rates. This sample was not exposed to ultra-high vacuum (UHV), which is required for TOF-SIMS analysis, until after AFM imaging, thus after crystallites had formed. This proves that UHV exposure is not the cause of crystallite formation. **a**, Ca surface concentration drops to background levels where **(b)** K is abundant, in discrete sites that are aligned along one of the cleavage directions and elongated along the other.

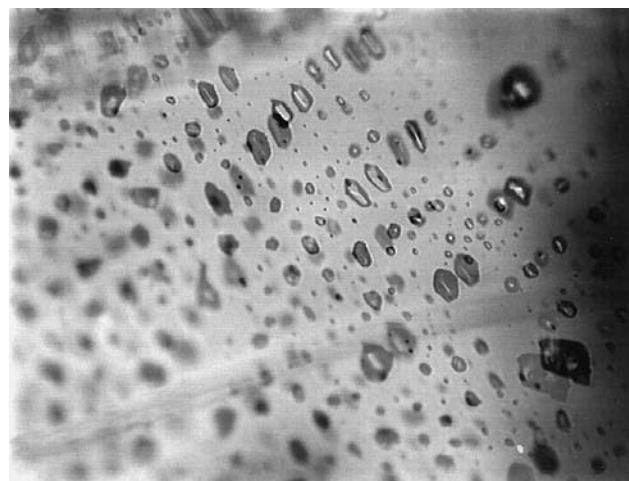


Figure 4 An optical microscope view through a calcite cleavage flake 1.5 mm thick. Fluid inclusions are elongated in one of the cleavage directions, and aligned along another. Heating-stage experiments show that the fluids are purely aqueous, containing Na^+ , K^+ , Cl^- , and Ca^{2+} or perhaps Mg^{2+} , with salinity in the range of about 0.9–1.6 eq. wt% NaCl. The unfocused line running across the lower part of the photograph is a large cleavage step on the bottom of the crystal. The photograph shows an area of calcite $620 \mu\text{m}$ wide; the smallest resolvable inclusions are $\sim 1 \mu\text{m}$ in diameter. We note that the limits of resolution for such a microscope are very much larger than for AFM and TOF-SIMS.

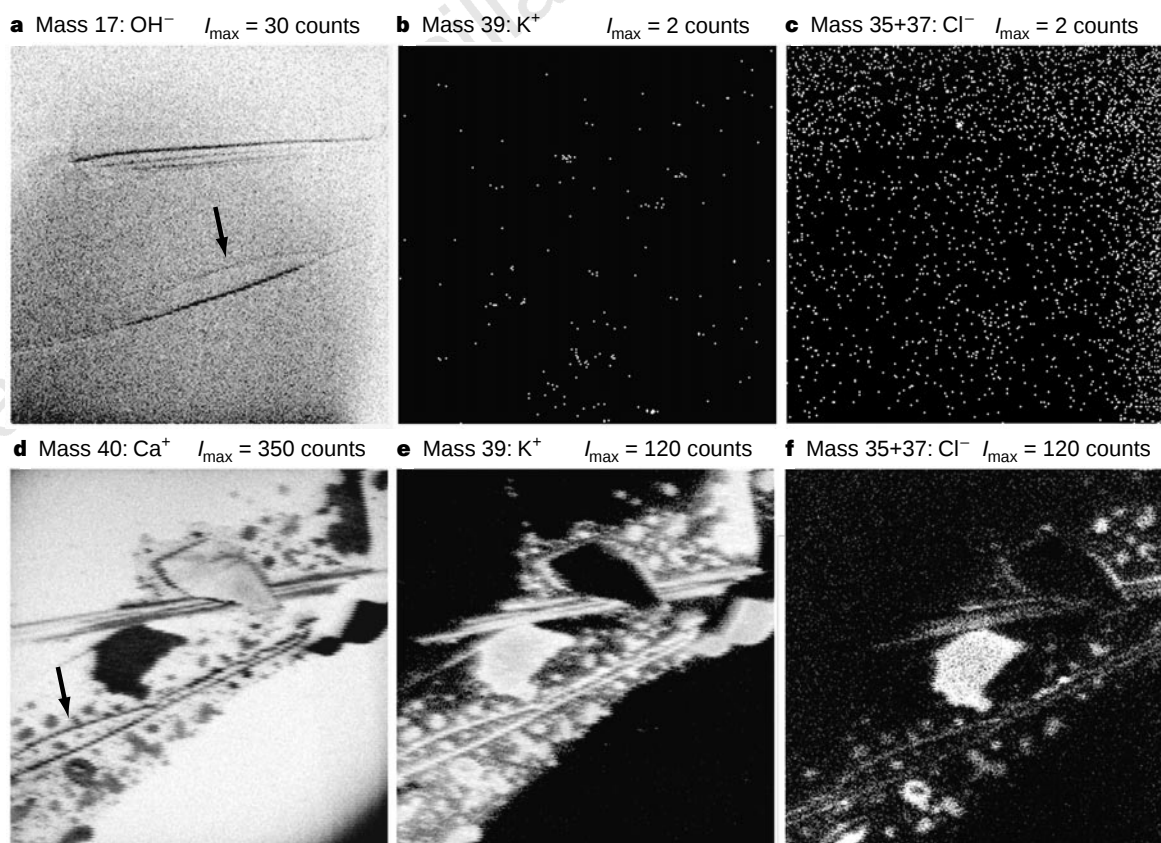


Figure 3 TOF-SIMS maps of relative elemental intensity from fresh (top row) and aged (bottom row) calcite surfaces. All images are $\sim 180 \mu\text{m}$ wide; slight variations in position and scaling (for the images in the bottom row) result from a difference in instrument function during collection of positive and negative ions. The arrow on images **a** and **d** indicates a 'marker' terrace edge, proving that the same site was found for reanalysis. " I_{max} " represents the maximum secondary ion intensity for the brightest areas; dark areas have low count rates. **a**, There is a high surface

concentration of hydroxide as expected, on the freshly cleaved sample, but surface concentrations of K (**b**) and Cl (**c**) are at background levels. **d–f**, After 5 months, enrichment of K and Cl parallels a decrease in surface concentration of Ca. In all experiments, chemical maps of O and hydroxide were similar to Ca; those of Na and F resembled K and Cl. On many samples, a few crystallites are more enriched in K^+ than Na^+ whereas in a few others, the reverse is true, suggesting some chemical segregation during formation.

consistently give rise to aged surfaces with many crystallites composed of monovalent species, but samples where fluid inclusions are sparse produce aged surfaces where such crystallites are lacking. Thus the accumulated evidence suggests that monovalent ions move out from fluid inclusions to form the observed surface features.

The remaining question is: what mechanism moves the material to the surface? In the absence of microfractures, we are left with the hypothesis of solid-state diffusion along linear defects. Exposure of freshly cleaved samples for 1 minute to distilled water at equilibrium with atmospheric CO₂ (pH ~ 5.6) typically results in the erosion of step edges and the nucleation on terraces of dissolution pits that are one atomic layer in depth (arrows, Fig. 1f). In contrast, on surfaces cleaved from samples where fluid inclusions are visible within the bulk and where crystallites form, exposure to water results in rows of very deep, narrow dissolution holes (Fig. 1f, black rhombs) aligned along one cleavage direction. We infer that these holes grow in response to higher free energy at defect sites. Linear dislocations could allow ions to move into, or out of, the solid at rates much faster than diffusion through ideally crystallized material.

To date, we know that Cd²⁺, Zn²⁺, Na⁺, K⁺ and Cl⁻ move in calcite at room temperature. We could find no reports describing similar behaviour for other ions, or for other ionic solids such as phosphates or chlorides. However, the observations presented here have important implications across several disciplines. As a contribution to fundamental mineral physics and chemistry, they show that free-energy inhibitions to ionic movement within a divalent ionic solid *can* be overcome, probably by hydration, and challenge us to define the mechanism. Although many calcite samples do provide Earth scientists with isotope ratios, trace-element distribution and fluid-inclusion compositions that are congruent with general geological relationships or with data from other minerals, solid-state diffusion of trace elements helps to explain those that do not. Evidence that some ions move through apparently perfectly crystalline calcite leads us to question what other ions and isotopes also move, especially in less ideal material. The white, opaque appearance of calcite is attributed to abundant defects and fluid inclusions¹⁶, suggesting higher defect density than in Iceland spar, and probably higher mobility. Our results warn that calcite may not provide the best samples for fluid-inclusion work or for isotope or trace-element studies.

In the field of environmental studies, predictions of the rate and extent of contaminant migration through aquifers come from models where mineral surfaces are treated as static and planar, with a finite number of surface sites. When these become saturated, uptake capacity of the porous medium is assumed to be exhausted. If surface sites are continually renewed or replenished by diffusion between surface and bulk, then the uptake or release capacity for calcite is currently underestimated. When crystallizing in water-treatment facilities, or in the high-pH conditions associated with fly-ash leachate or the breakdown of cement repositories, calcite incorporates toxic metals and radionuclides. If trace components move into or out of the bulk, then their concentration in the contacting solution is altered, even in waters at equilibrium with calcite. □

Received 4 March; accepted 28 August 1998.

1. Wada, H. Microscale isotopic zoning in calcite and graphite crystals in marble. *Nature* **331**, 61–63 (1988).
2. Farver, J. R. Oxygen self-diffusion in calcite: Dependence on temperature and water fugacity. *Earth Planet. Sci. Lett.* **121**, 575–587 (1994).
3. Cherniak, D. J. An experimental study of strontium and lead diffusion in calcite, and implications for carbonate diagenesis and metamorphism. *Geochim. Cosmochim. Acta* **61**, 4173–4179 (1997).
4. Lahann, R. W. & Siebert, R. M. A kinetic model for distribution coefficients and application to Mg-calcites. *Geochim. Cosmochim. Acta* **46**, 2229–2237 (1982).
5. Davis, J. A., Fuller, C. C. & Cook, A. D. A model for trace metal sorption processes at the calcite surface: Adsorption of Cd²⁺ and subsequent solid solution formation. *Geochim. Cosmochim. Acta* **51**, 1477–1490 (1987).
6. Lahav, N. & Bolt, G. H. Self-diffusion of ⁴⁵Ca into certain carbonates. *Soil Sci.* **97**, 293–299 (1964).
7. Stipp, S. L., Hochella, M. F. Jr, Parks, G. A. & Leckie, J. O. Cd²⁺ uptake by calcite, solid-state diffusion, and the formation of solid-solution: Interface processes observed with near-surface sensitive techniques (XPS, LEED, and AES). *Geochim. Cosmochim. Acta* **56**, 1941–1954 (1992).
8. Stipp, S. L. S. Understanding interface processes and their role in the mobility of contaminants in the geosphere: The use of surface sensitive techniques. *Eclogae Geol. Helv.* **87**, 335–355 (1994).
9. Reeder, R. *Carbonates: Mineralogy and Chemistry*. (Reviews in Mineralogy Vol. 11, Mineralogical Society of America, Washington DC, 1983).

10. Plummer, L. N. & Busenberg, E. The solubilities of calcite, aragonite, and vaterite in CO₂-H₂O solutions between 0° and 90 °C, and an evaluation of the aqueous model for the system CaCO₃-CO₂-H₂O. *Geochim. Cosmochim. Acta* **46**, 1011–1040 (1982).
11. Eggleston, C. M. in *Scanning Probe Microscopy of Clay Minerals* (eds Nagy, K. L. & Blum, A. E.) 1–90 (Vol. 7, Workshop Lectures, Clay Mineral. Soc., Boulder, CO, 1994).
12. Stipp, S. L. S., Gutmannsbauer, W. & Lehmann, T. The dynamic nature of calcite surfaces in air. *Am. Mineral.* **81**, 1–8 (1996).
13. Chiarello, R. P., Wogelius, R. A. & Sturchio, N. C. In-situ synchrotron X-ray reflectivity measurements at the calcite-water interface. *Geochim. Cosmochim. Acta* **57**, 4103–4110 (1993).
14. Benninghoven, A., Hagenhoff, B. & Neihus, E. Surface MS-probing of real-world samples. *Anal. Chem.* **65**, 630A–640A (1993).
15. Stipp, S. L. S., Kulik, A. J., Franzreb, K., Benoit, W. & Mathieu, H. J. A combination of SEM and TOF-SIMS imaging for observing local inhomogeneities in morphology and composition. *Surf. Interface Anal.* **25**, 959–965 (1997).
16. Roedder, E. *Fluid Inclusions*. (Reviews in Mineralogy Vol. 12, Mineralogical Society of America, Washington DC, 1984).

Acknowledgements. We thank W. Benoit, H.-R. Pfeifer and J.-C. Védry for support during the early stages of the work; J. Ronsbo and E. Makovsky for discussions; U. Hoffmann for help with collection of Fig. 1f; P. Shi, D. Léonard and Y. Chevolut for instrument support; and B. Svane Nielsen, R. Bromley, B. Buchardt, H. J. Hansen, R. Yund and G. Jenkin for comments. This work is dedicated to George A. Parks. Partial funding was provided by the Danish Research Council.

Correspondence and requests for materials should be addressed to S.L.S.S. (e-mail: stipp@geo.geol.ku.dk).

Large variations in vent fluid CO₂/³He ratios signal rapid changes in magma chemistry at Loihi seamount, Hawaii

D. R. Hilton*, G. M. McMurtry† & F. Goff‡

* Isotope Laboratory, Geosciences Research Division, Scripps Institution of Oceanography, La Jolla, California 92093-0220, USA

† School of Ocean & Earth Science & Technology, University of Hawaii, Honolulu, Hawaii 96822, USA

‡ Earth & Environmental Sciences Division, Los Alamos National Laboratory, Los Alamos, New Mexico 87545, USA

Loihi seamount, an active submarine volcano situated about 30 km south of the island of Hawaii, is the youngest manifestation of the hotspot responsible for the Emperor–Hawaiian seamount chain and Hawaiian islands. This seamount has been the focus of numerous studies characterizing the geophysical, geochemical and biological features of an active intraplate volcano^{1–14}. In July–August 1996, Loihi seamount experienced the most intense period of seismic activity yet recorded for any Hawaiian volcano¹. Within two months of the ‘seismic crisis’, summit and flank hydrothermal vent fluids were collected using a manned submersible. Here we report data from these samples that indicate large and systematic changes in the CO₂/³He ratios of the vent fluids compared to pre-seismic-crisis values^{2,3}. These changes are consistent with an abrupt transition from alkalic to tholeiitic basaltic magma having supplied volatiles to the vents. This rapid change in magma chemistry has been discernible only through CO₂/³He monitoring, and suggests that the anticipated evolution of the Hawaiian plume to a phase of shield-building tholeiitic magmatism is highly episodic at Loihi and not yet complete.

The summit of Loihi seamount is located at a depth of ~1,000 m on the southern flanks of Mauna Loa and Kilauea volcanoes of the island of Hawaii (Fig. 1). Following the seismic swarms of 1996, a Rapid Response Cruise (RRC) and two previously scheduled LONO (Loihi Oceanographic Neotectonic Observations) cruises were dispatched to Loihi to record features and consequences of the seismicity¹. The site at Pele’s vents^{2–4} was obliterated by the formation of a new pit crater; however, a number of new vent fields with significantly higher temperatures (up to 77 °C) appeared instead (Fig. 1). In Table 1, we document the helium and carbon (and heat) characteristics of three of these newly formed vent fields: Lohiau,

located on the summit, and Pohaku and Naha located on the South rift.

Marked changes in $^3\text{He}/^4\text{He}$ and $\text{CO}_2/^3\text{He}$ ratios are apparent from the 1996 sampling programme when compared with the 1992–93 values and relative to initial reports of these parameters following the discovery of Pele's vents in 1987 (Fig. 2). The $^3\text{He}/^4\text{He}$ ratios in 1996 (values up to $25.5R_A$ where R_A is the $^3\text{He}/^4\text{He}$ of air) are discernibly lower than 1992–93 values (maximum $(27.0\text{--}27.9)R_A$), indicating a small but perceptible change in the proportion of hotspot helium (characterized as $>30R_A$) mixing with either MORB-like ($8R_A$) and/or radiogenic helium ($0.05R_A$) (ref. 15 details the $^3\text{He}/^4\text{He}$ characteristics of these components; MORB, mid-ocean ridge basalt). Significantly, there is not a helium-isotope difference (within analytical error) between the Lohiau summit vents ($25.5R_A$) and the Naha and Pohaku, South rift, vents ($24.7R_A$). In addition, these 1996 fluid $^3\text{He}/^4\text{He}$ ratios match almost exactly the $^3\text{He}/^4\text{He}$ ratio measured in a tholeiitic basalt ($25.0R_A$)⁵ erupted in early 1996 immediately before the seismic crisis. We conclude that helium sampled by the summit and flank fluids following the seismic crisis is derived from the same magmatic source as the basalt erupted earlier in the year. The wide range in $^3\text{He}/^4\text{He}$ ratios ($(23.0\text{--}35)R_A$) now reported for hydrothermal fluids (this study and refs 2 and 3) and lavas of Loihi seamount (for example, refs 6–8) indicates that individual magma batches may record different mixing histories between mantle and lithospheric sources; this is because magma ageing effects are difficult to reconcile with evidence of relatively short magma transit times through the Hawaiian lithosphere¹⁶.

The most dramatic variation of a geochemical parameter observed at Loihi, both in a temporal and spatial sense (that is, between the summit and flank localities), is for the $\text{CO}_2/^3\text{He}$ ratios (Fig. 2). The summit (Lohiau) vents have $\text{CO}_2/^3\text{He}$ ratios consistently lower than 10^9 —irrespective of type of sampler employed, individual vent site sampled and date of sampling (Table 1). The vent-fluid ratios at the summit overlap with the 1996 tholeiitic basalt ($\text{CO}_2/^3\text{He} = (2\text{--}5) \times 10^8$; ref. 5), again indicating that both sampling media access these volatiles from a common provenance without bias. To date, such low values are unprecedented in studies of hydrothermal vent fluids. The flank fluids (Naha and Pohaku) have higher values, $(2\text{--}4) \times 10^9$, which are considered typical of MORB¹⁷ and coincide with ratios observed in 1987 at Loihi². Since the last period of sampling (1992–93)³, therefore, there has been a spectacular decrease—from values in the range $(8\text{--}27) \times 10^9$ —in the summit $\text{CO}_2/^3\text{He}$ ratios. Can these order-of-magnitude changes be related to magma movement and degassing associated with the 1996 seismic crisis?

There are three factors relevant to the range of $\text{CO}_2/^3\text{He}$ ratios observed at Loihi seamount: (1) possible variations in the initial or starting $\text{CO}_2/^3\text{He}$ ratio of the magma; (2) the fractionation factor ($\alpha_{\text{CO}_2\text{--He}}$) describing how carbon and helium fractionate as they partition between a vapour phase and the melt (it is the vapour phase which is sampled by vesicles in a lava, and which presumably also supplies the hydrothermal vents); and (3) the style of degassing, that is, open system (or Rayleigh) distillation whereby vapour is removed from contact with the melt as it is formed, or closed-system (or batch) distillation in which it stays in contact with the melt. We can dismiss the first option with a high degree of confidence for two reasons. First, the two most plausible contributors to Hawaiian magma sources (hotspot and MORB–mantle) have similar $\text{CO}_2/^3\text{He}$ ratios ($\sim(2\text{--}5) \times 10^9$; refs 17–19), thereby precluding variations in mixing proportions of the potential end-members as a cause of the low $\text{CO}_2/^3\text{He}$ values. Second, overlapping $^3\text{He}/^4\text{He}$ ratios at the summit and South rift (flank) localities in 1996 indicate that both are likely to be tapping the same volatile source.

The fractionation factor α , defined as $(\text{CO}_2/^3\text{He})_{\text{vap}}/(\text{CO}_2/^3\text{He})_{\text{melt}}$, is given by the inverse ratio of their solubilities (S) in the parental melt. For a tholeiitic melt ($\text{SiO}_2 \sim 48\text{--}49\text{ wt\%}$), such as that erupted

in 1996⁵, the solubilities of CO_2 and He are expected to be $25.5 \times 10^{-5} \text{ cm}^3 \text{ STP g}^{-1}$ (refs 20, 21) and $(56\text{--}64) \times 10^{-5} \text{ cm}^3 \text{ STP g}^{-1}$ (refs 22, 23) respectively; therefore, $\alpha (= S_{\text{He}}/S_{\text{CO}_2})$ is estimated at a mean value of 2.35. If the starting $\text{CO}_2/^3\text{He}$ ratio of hotspot volatiles before gas loss is $\sim 2 \times 10^9$, then vesiculation will lead to an increase in the $\text{CO}_2/^3\text{He}$ ratio in the vesicle phase and corresponding decrease in residual volatiles dissolved in the melt. Irrespective of the mode of degassing, the maximum $\text{CO}_2/^3\text{He}$ of the vesicle phase is constrained to be 2.35 times the starting value (4.7×10^9); however, the $\text{CO}_2/^3\text{He}$ ratio of residual volatiles in the melt is highly dependent upon the mode of degassing (batch or Rayleigh distillation). In the former case, the minimum $\text{CO}_2/^3\text{He}$ ratio attainable is 8.5×10^8 for close to 100% transfer of volatiles into the vapour phase (starting value divided by 2.35). In the latter mode, however, any $\text{CO}_2/^3\text{He}$ ratio is possible depending upon the fraction of volatiles lost. Using the Rayleigh distillation equation $(\text{CO}_2/^3\text{He})_{\text{res}} = (\text{CO}_2/^3\text{He})_{\text{initial}} F \exp[(\alpha - 1)/\alpha]$ where F is the fraction of residual (res) CO_2 in the magma, we calculate that loss of 77% of the CO_2 will produce the above $\text{CO}_2/^3\text{He}$ ratio of 8.5×10^8 , and a loss of 88% will give 5.9×10^8 (the lowest recorded value for 1996; Table 1). Sampling volatiles remaining after a degassing event is a plausible explanation for the $\text{CO}_2/^3\text{He}$ ratios of Lohiau vent fluids; it is consistent with the significantly lower CO_2 concentrations in these fluids (maximum, $49.7 \text{ mmol kg}^{-1}$; Table 1) com-

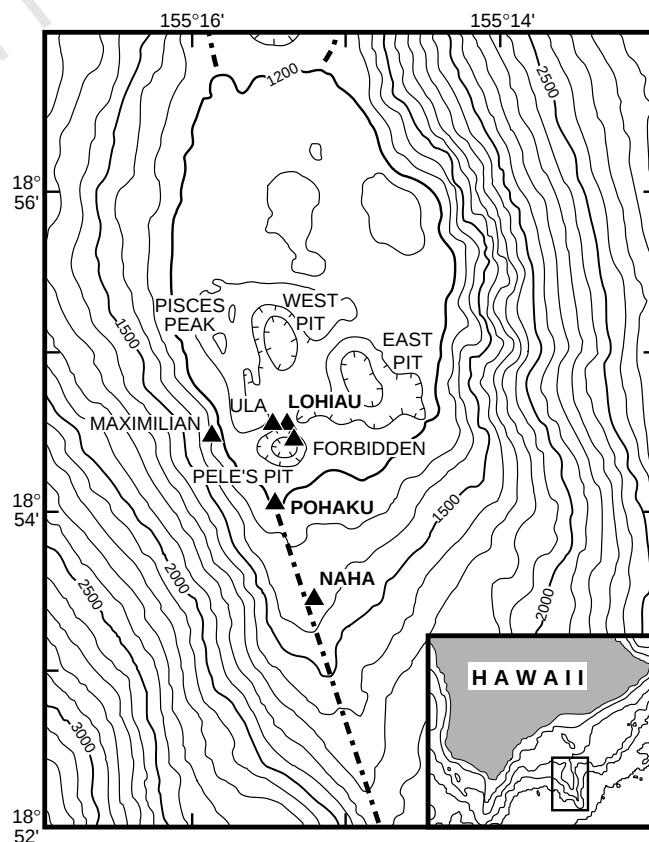


Figure 1 Bathymetry of Loihi seamount following the seismic crisis of July–August 1996. The summit platform of the volcano is defined by the 1,200-m contour, and filled triangles designate the hydrothermal vent fields. The Lohiau and Ula vent fields are located in altered pillow basalt outcrops on the steep northern wall of the new Pele's pit crater. Their position and that of the Forbidden vent field at the basal northeast wall of the pit crater suggest circulation within a ring fault system along the margins of the crater. The Pohaku and Naha vent fields are located within fissured outcrops of relatively fresh pillow basalts along the axis of the South rift. Maximilian vent field lies within basalt outcrops on the steep, talus-covered western flank of Loihi. Vent fields sampled during the LONO expedition (this work) are shown in bold type.

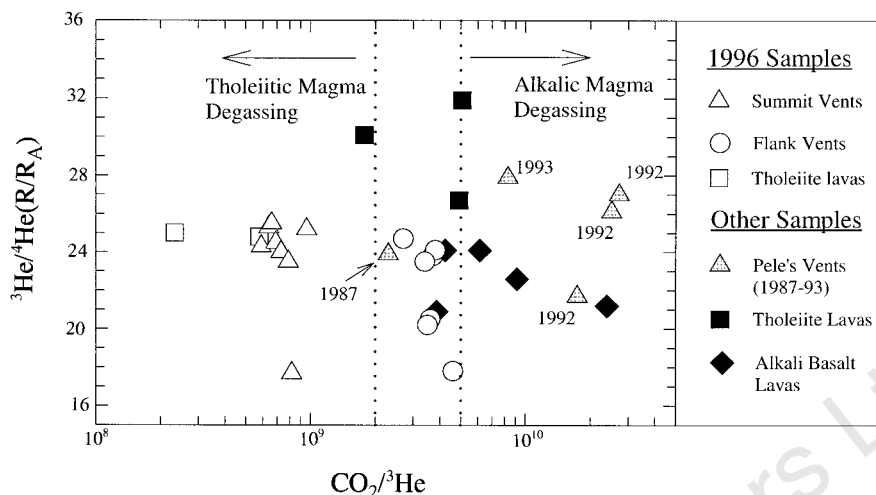


Figure 2 Helium isotope ratios versus CO_2/He ratios. Shown are Loihi seamount $^3\text{He}/^4\text{He}$ ratios (R/R_A notation) versus CO_2/He ratios for hydrothermal fluids and tholeiitic basalts collected in 1996 (open symbols; this work and ref. 5), vent fluids from the recently-destroyed Pele's vents (grey symbols; refs 2, 3), and tholeiitic and alkalic lavas of unknown age (filled symbols; refs 7 and 30). We note that vesicle gas (released by crushing samples *in vacuo*) is plotted for the 1996 tholeiite samples—so CO_2/He ratios may be directly comparable to hydrothermal vents which also sample exsolved gas—whereas the “Other Samples” tholeiite and alkali basalts CO_2/He values are calculated using (where possible) total helium abundances (vesicle plus a minor dissolved component)⁷ and magmatic carbon abundances (that is, carbon released above 800 °C by stepwise melting which therefore includes vesicle and dissolved carbon)³⁰. Although the ‘magmatic’ CO_2/He ratio is plotted, the high vesicularities of all these lavas (2–40%)⁷

indicate that this value will approximate that of the vesicles (see Fig. 5 in ref. 22). Therefore, post-seismic-crisis 1996 hydrothermal fluids (this work) overlap in CO_2/He ratios with basalts of tholeiitic composition whereas pre-seismic-crisis 1992/3 fluids coincide with lavas of alkalic affinity. We note that there is significant overlap in fluid and lava CO_2/He values in the range $(2\text{--}5) \times 10^9$; this range most probably represents the initial or ‘starting’ CO_2/He ratio at Loihi seamount. Although the anticipated evolutionary trajectories of CO_2/He ratios of residual volatiles after degassing are given for the two magma compositions, the absolute shift in CO_2/He ratio from the initial value following a degassing event(s) will depend upon the exact magma composition (and hence absolute value of α), mode of degassing (Rayleigh versus closed-system) and extent of degassing (see text). Degassing is not expected to affect the $^3\text{He}/^4\text{He}$ ratio of lavas or vent fluids.

Table 1 LONO cruise—helium, carbon and heat relationships of Loihi seamount vent fluids

Dive (sampler*)	Date of sampling	SiO_2 ($\mu\text{mol kg}^{-1}$)	Temp ^t (°C)	R/R_A	(He/Ne) _{sample} / (He/Ne) _{air}	R_C/R_A †	[CO ₂] (mmol kg ⁻¹)	CO_2/He	$^3\text{He}/\text{heat}$ § ($\times 10^{-12} \text{ cm}^3 \text{ STP cal}^{-1}$)
Summit (Lohiau) vent field									
310 no. 1	1 Oct. 96	1,730	NA	25.1	97.8	25.3	24.6	6.4×10^8	14.4
310 no. 2	1 Oct. 96	466	NA	24.3	77.4	24.3	4.79	5.9×10^8	9.1
311 no. 3	2 Oct. 96	2,396	77.0	25.1	396	25.2	49.7	9.6×10^8	14.3
311 (LaB)	2 Oct. 96	1,212	77.0	23.8	27.3	24.5	5.59	6.9×10^8	4.2
314 no. 1	6 Oct. 96	975	51.6	23.9	191	24.0	12.7	7.3×10^8	10.8
316 no. 1	8 Oct. 96	969	56.0	17.1	19.4	17.7	11.0	8.2×10^8	8.4
316 no. 2	8 Oct. 96	1,340	56.0	25.3	91.8	25.5	18.8	6.6×10^8	13.4
316 (LaB)	8 Oct. 96	1,440	52.9	23.4	134	23.5	17.9	7.9×10^8	10.0
South rift (Naha vents)									
312 no. 1	3 Oct. 96	226	11.2	19.0	10.1	20.5	1.01	3.6×10^9	1.53
312 no. 2	3 Oct. 96	190	11.2	14.9	2.89	20.2	0.42	3.5×10^9	0.77
313 no. 1	4 Oct. 96	659	22.8	23.2	29.2	23.8	8.28	3.7×10^9	4.25
313 no. 2	4 Oct. 96	1,638	22.8	23.5	32.4	24.1	20.2	3.8×10^9	4.08
313 (LaB)	4 Oct. 96	737	22.8	22.7	23.6	23.5	6.82	3.4×10^9	3.45
South rift (Pohaku vents)									
315 no. 1	7 Oct. 96	1,240	17.2	17.2	24.1	17.8	37.7	4.6×10^9	8.31
315 no. 2	7 Oct. 96	602	17.2	24.6	115	24.7	15.4	2.7×10^9	11.84
315 (LaB)	7 Oct. 96	541	13.7	11.5	9.82	12.4	4.65	2.9×10^9	3.64
Summit (Pele's vents)¶									
93LS-3c (LaB)	7 Sept. 93	914	18.5	27.2	28.5	27.9	99.5	8.3×10^9	18.6

LONO expedition information, including participant list, daily reports and sample site descriptions are available at <http://www.soest.hawaii.edu/GG/HCV/loihi.html>

* LaB = La Bomba sampler (described in ref. 3); other samplers (‘Major’ bottles) described in ref. 28.

† $R_C/R_A = \{(R/R_A \times X) - 1\} / (X - 1)$ where $X = \{(\text{He/Ne})_{\text{sample}} / (\text{He/Ne})_{\text{air}}\} \times \{\beta_{\text{He}} / \beta_{\text{He}}\}$ and β is the Bunsen solubility coefficient. $\{\beta_{\text{He}} / \beta_{\text{He}}\} = 1.261$ (ref. 29) at 4 °C and 35‰ salinity. Analytical uncertainty on R_C/R_A is $\pm 3\%$ (and that on CO_2/He is $\pm 5\%$) which reflects daily variations in the MM standard at the Isotope Laboratory. Blanks were negligible compared to sample sizes (see ref. 6 for details).

‡ Temperature measured with probe immediately prior to sampling. NA, not available.

§ Calculated using regression fits: SiO_2 ($\mu\text{mol kg}^{-1}$) = $31.67(^\circ\text{C}) - 296$; $r^2 = 0.84$, $n = 7$ (Lohiau vents) or SiO_2 ($\mu\text{mol kg}^{-1}$) = $56.27(^\circ\text{C}) - 232$; $r^2 = 0.53$, $n = 9$ (South rift vents) adopting above (and other LONO cruise) data. Use of SiO_2 –temperature relationships for various vent fields gives a more reliable estimate of individual fluid enthalpies as fluids may be sampled at slightly different points from where temperature measurements were made (see discussion in refs 3, 4). It also allows calculation of $^3\text{He}/\text{heat}$ values for dive nos 310 based on measured SiO_2 . Data to be discussed elsewhere (G.M. McM. *et al.*, manuscript in preparation).

¶ Previously unpublished sample from 1993. Vent site no longer in existence.

pared with values as high as 321 and 190 mmol kg⁻¹ found at the summit in 1987 and 1992, respectively. If the Naha and Pohaku vents are supplied by volatiles from the same magma source (which is plausible given the overlap in ³He/⁴He ratios), then for a starting ratio of 2 × 10⁹ they must sample vapour phase CO₂ which is complementary to residual CO₂ sampled at the summit. Alternatively, if the starting CO₂/³He was slightly higher (for example, 5 × 10⁹; see Fig. 2), both summit and flanks could be sampling residual volatiles but the degree of degassing experienced by the magma batch supplying Lohiau vents would have to be greater than that supplying the flank localities.

If the initial CO₂/³He ratio of Loihi magma lies in the range (2–5) × 10⁹, then it is difficult to account for CO₂/³He ratios of (8–27) × 10⁹—as observed in 1992 and 1993—using the scheme described above. A possible explanation lies with the solubility behaviour of CO₂ and He in magmas of different chemistry. In low-SiO₂ alkalic basalts, characteristic of the pre-shield-building stage of Hawaiian volcanism, the solubility of CO₂ is significantly higher than in tholeiites²⁴ which represent the more mature and voluminous shield-building phase: the opposite is true of helium whose solubility decreases with decreasing SiO₂ (ref. 23). Both effects act to reduce the value of the fractionation factor ($\alpha_{\text{CO}_2\text{-He}}$) in magmas of alkalic versus tholeiitic chemistries. Using the solubility parametrizations of Dixon (ref. 24) and Lux (ref. 23), $\alpha_{\text{CO}_2\text{-He}}$ will become less than unity at SiO₂ contents in the range 45–46 wt% (we assume that water influences He solubility in a similar fashion to CO₂ solubility²⁰ so that $\alpha_{\text{CO}_2\text{-He}}$ is independent of dissolved water content). At such values, the evolution of CO₂/³He as a function of degassing history is likely to be the opposite of that described above: volatiles residual after a degassing event will have higher CO₂/³He. The magnitude of this ratio will depend again on the value of α and mode (batch versus Rayleigh) and extent (F) of degassing. Degassing of an alkalic magma may explain why CO₂/³He ratios reached maximum values of 27.3 × 10⁹ in 1992 and were still high in 1993 (CO₂/³He = 8.3 × 10⁹).

The variations in the CO₂/³He ratios in Loihi vent fluids place strong constraints on the type of magma degassing at Loihi. Previous work has documented that the summit of Loihi seamount is dominated by tholeiitic lavas with only relatively rare occurrences of alkalic basalts^{10,11}. This observation is consistent with increased degrees of partial melting and/or shallower depths of initiation of melting of the source region as the Pacific plate moves Loihi over the centre of the Hawaiian plume^{9–11,25,26}; it has led to the notion that the transition from the pre-shield-building alkalic phase to the main shield-building tholeiitic phase at Loihi is essentially complete^{10,11}. The duration of the transition has been estimated at 17–40 kyr, similar to that of other Hawaiian volcanoes²⁷. The high CO₂/³He ratios (> 8 × 10⁹) observed in the fluids over the past decade confirm the recent presence of alkalic magma in the summit region—a conclusion not possible to reach by other means—and indicate that the transition phase is still continuing at Loihi. Furthermore, our data suggest that transitions between alkalic and tholeiitic volcanism can occur rapidly, and we estimate a maximum of 2–3 years for the most recent transition based on the interval between our 1993 sampling programme and the initiation of the seismic crisis. Whether the low CO₂/³He ratios of 1996 (which are consistent with degassing of tholeiitic magmas) herald the final transition to the shield-building phase, or are a temporary change in degree or depth of partial melting, underscores the importance of continued monitoring of this useful vent-fluid parameter. □

Received 15 January; accepted 19 August 1998.

1. Duennebie, F. K. et al. Researchers rapidly respond to submarine activity at Loihi volcano, Hawaii. *Eos* **78**, 229–233 (1997).
2. Craig, H., Welhan, J. A. & Hilton, D. R. Hydrothermal vents in Loihi caldera: Alvin results. *Eos* **68**, 1553 (1987).
3. Sedwick, P., McMurtry, G. M., Hilton, D. R. & Goff, F. Carbon dioxide and helium in hydrothermal fluids from Loihi Seamount, Hawaii, USA: Temporal variability and implications for the release of mantle volatiles. *Geochim. Cosmochim. Acta* **58**, 1219–1227 (1994).

4. Sedwick, P., McMurtry, G. M. & Macdougall, J. D. Chemistry of hydrothermal solutions from Pele's Vents, Loihi Seamount, Hawaii. *Geochim. Cosmochim. Acta* **56**, 3643–3667 (1992).
5. Garcia, M. O. et al. Petrology and geochronology of basalt breccia from the 1996 earthquake swarm of Loihi Seamount, Hawaii: magmatic history of its 1996 eruption. *Bull. Volcanol.* **59**, 577–592 (1998).
6. Rison, W. & Craig, H. Helium isotopes and mantle volatiles in Loihi Seamount and Hawaiian Island basalts and xenoliths. *Earth Planet. Sci. Lett.* **66**, 407–426 (1983).
7. Kurz, M. D., Jenkins, W. J., Hart, S. R. & Clague, D. A. Helium isotopic variations in volcanic rocks from Loihi Seamount and the Island of Hawaii. *Earth Planet. Sci. Lett.* **66**, 388–406 (1983).
8. Valbrach, P. J., Staudigel, H., Honda, M., McDougall, I. & Davies, G. R. Isotopic tracing of volcanic source regions from Hawaii: decoupling of gaseous from lithophile magma components. *Earth Planet. Sci. Lett.* **144**, 185–198 (1996).
9. Moore, J. G., Clague, D. A. & Normark, W. R. Diverse basalt types from Loihi Seamount. *Geology* **10**, 88–92 (1982).
10. Garcia, M. O., Jorgenson, B., Mahoney, J. J., Ito, E. & Irving, A. J. An evaluation of temporal geochemical evolution of Loihi summit lavas: Results from Alvin submersible dives. *J. Geophys. Res.* **98**, 537–550 (1993).
11. Garcia, M. O., Foss, D. J. P., West, W. B. & Mahoney, J. J. Geochemical and isotopic evolution of Loihi volcano, Hawaii. *J. Petrol.* **26**, 1647–1674 (1995).
12. Malahoff, A., McMurtry, G. M., Wiltshire, J. C. & Yeh, H.-W. Geology and chemistry of hydrothermal deposits from active submarine volcano Loihi, Hawaii. *Nature* **298**, 234–239 (1982).
13. Karl, D., McMurtry, G. M., Malahoff, A. & Garcia, M. Loihi Seamount, Hawaii: A mid-plate volcano with a distinctive hydrothermal system. *Nature* **335**, 532–535 (1988).
14. Moyer, C. L., Dobbs, F. C. & Karl, D. M. Estimation of diversity and community structure through restriction fragment length polymorphism distribution analysis of bacterial 16S rRNA genes from a microbial mat at an active, hydrothermal vent system, Loihi Seamount, Hawaii. *Appl. Environ. Microbiol.* **60**, 871–879 (1994).
15. Craig, H. & Lupton, J. E. in *The Sea* Vol. 7 (ed. Emiliani, C.) 391–428 (Wiley, New York, 1981).
16. Cohen, A. S., O'Nions, R. K. & Kurz, M. D. Chemical and isotopic variations in Mauna Loa tholeiites. *Earth Planet. Sci. Lett.* **143**, 111–124 (1996).
17. Marty, B. & Jambon, A. C³He in volatile fluxes from the solid Earth: Implications for carbon geodynamics. *Earth Planet. Sci. Lett.* **83**, 16–26 (1987).
18. Trull, T., Nadeau, S., Pineau, F., Polve, M. & Javoy, M. C-He systematics in hotspot xenoliths: Implications for mantle carbon contents and carbon recycling. *Earth Planet. Sci. Lett.* **118**, 43–64 (1993).
19. Hilton, D. R., McMurtry, G. M. & Kreulen, R. Evidence for extensive degassing of the Hawaiian mantle plume from helium-carbon relationships at Kilauea volcano. *Geophys. Res. Lett.* **24**, 3065–3068 (1997).
20. Dixon, J. E., Stolper, E. M. & Holloway, J. R. An experimental study of water and carbon dioxide solubilities in mid-ocean ridge basaltic liquids. Part I: Calibration and solubility models. *J. Petrol.* **36**, 1607–1631 (1995).
21. Jendzejewski, N., Trull, T. W., Pineau, F. & Javoy, M. Carbon solubility in mid-ocean ridge basaltic melt at low pressures (250–1950 bar). *Chem. Geol.* **138**, 81–92 (1997).
22. Jambon, A., Weber, H. & Braun, O. Solubility of He, Ne, Ar, Kr and Xe in a basalt melt in the range 1250–1600 °C. Geochemical implications. *Geochim. Cosmochim. Acta* **50**, 401–408 (1986).
23. Lux, G. The behavior of noble gases in silicate liquids: Solution, diffusion, bubbles and surface effects, with applications to natural samples. *Geochim. Cosmochim. Acta* **51**, 1549–1560 (1987).
24. Dixon, J. E. Degassing of alkalic basalts. *Am. Mineral.* **82**, 368–378 (1997).
25. Frey, F. A. & Clague, D. A. Geochemistry of diverse basalt types from Loihi Seamount, Hawaii: petrogenetic implications. *Earth Planet. Sci. Lett.* **66**, 337–355 (1983).
26. Frey, F. A. & Rhodes, J. M. Intersheet geochemical differences among Hawaiian volcanoes: implications for source compositions, melting process and magma ascent paths. *Phil. Trans. R. Soc. Lond. A* **342**, 121–136 (1993).
27. Guillou, H., Garcia, M. O. & Turpin, L. Unspiked K-Ar dating of young volcanic rocks from the Loihi and Pitcairn seamounts. *J. Volcanol. Geotherm. Res.* **78**, 239–250 (1997).
28. Von Damm, K. L. et al. Chemistry of submarine hydrothermal solutions at 21 degrees N, East Pacific Rise. *Geochim. Cosmochim. Acta* **49**, 2197–2220 (1985).
29. Ozima, M. & Podosek, F. A. *Noble Gas Geochemistry* (Cambridge Univ. Press, 1983).
30. Exley, R. A., Matthey, D. P., Clague, D. A. & Pillinger, C. T. Carbon isotope systematics of a mantle "hotspot": a comparison of Loihi Seamount and MORB glasses. *Earth Planet. Sci. Lett.* **78**, 189–199 (1986).

Acknowledgements. We thank F. Duennebie and J. R. Smith for Fig. 1, and B. Eakins, M. Cremer and D. Counce for laboratory assistance. We also thank T. Kerby and the HURL operations staff for help at sea. Discussions with M. O. Garcia and C. G. Macpherson proved very helpful. This work was funded by the NOAA Sea Grant College and National Undersea Research Programs (to G.M.M.) and SIO "start-up" funds (to D.R.H.).

Correspondence and requests for materials should be addressed to D.R.H. (e-mail: drhilton@ucsd.edu).

Neuronal synchrony does not represent texture segregation

Victor A. F. Lamme & Henk Spekreijse

Graduate School Neurosciences Amsterdam, AMC, Department of Medical Physics, University of Amsterdam, and The Netherlands Ophthalmic Research Institute, PO Box 12141, 1100 AC Amsterdam, The Netherlands

The visual environment is perceived as an organized whole of objects and their surroundings. In many visual cortical areas, however, neurons are typically activated when a stimulus is presented over a very limited portion of the visual field, the receptive field of that neuron^{1–4}. To bridge the gap between this piecemeal neuronal analysis and our global visual percepts, it has been postulated that neurons representing elements of the same

object fire in synchrony to represent the perceptual organization of a scene⁵⁻¹⁰. Experiments with stimuli such as moving bars or gratings have provided evidence for this hypothesis¹¹⁻¹⁶. We have further tested this by presenting monkeys with various textured scenes consisting of a figure on a background, and recorded neuronal activity in the primary visual cortex (area V1). Our results show no systematic relationship between the synchrony of firing of pairs of neurons and the perceptual organization of the scene. Instead, pairs of recording sites representing elements of the same figure most commonly showed equal amounts of synchrony between them as did pairs of which one site represented the figure and the other the background. We conclude that synchrony in V1 does not reflect the binding of features that leads to texture segregation.

We used 90° and 20° differences in orientation of line segments

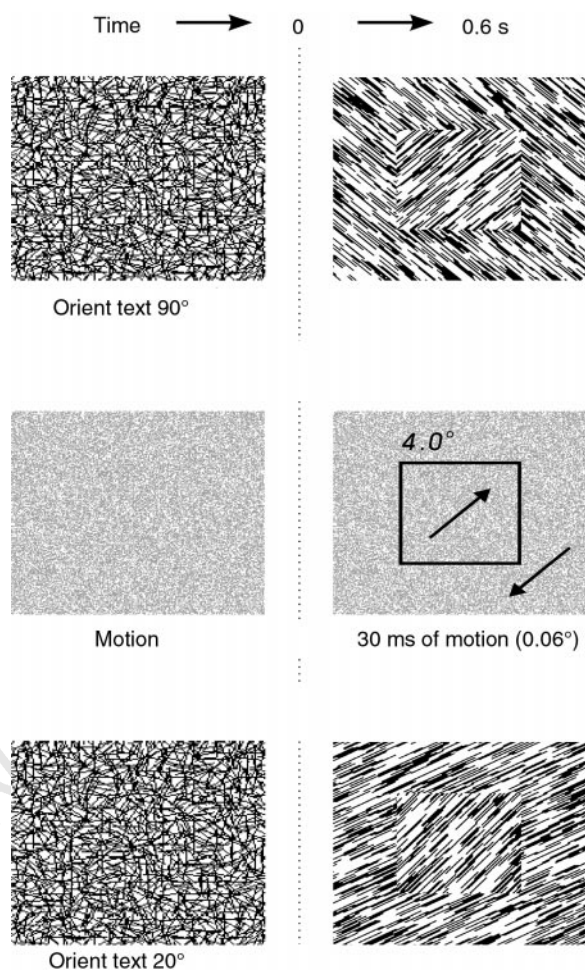


Figure 1 Illustrations (not replications) of the three types of figure-ground display used in this study. The screen size (28° × 21°) was larger than suggested by these drawings. 'Orient text 90°', 'Orient text 20°': at time zero, a texture of randomly oriented and positioned line segments (left) is instantly replaced by a figure-ground texture with a 90° or 20° difference in orientation (right). 'Motion': at time zero, previously stationary random dots inside and outside a 4° square region start to move in opposite directions. The motion 'impulse' lasts 30 ms while dots move over a distance of two pixels (0.06°). The figure-ground displays were on for 600 ms, while the animal maintained fixation.

between figure and ground and the motion of random dots in opposite directions as cues to segregate a textured square figure from a textured background (Fig. 1). Neuronal activity was recorded simultaneously from 16 implanted microwire electrodes on each monkey. The figure (4°) was much larger than the aggregate receptive field sizes (typically 0.5°) of the multi-unit activity that could be recorded at each retinotopic site in V1. The figure was presented at various positions relative to the receptive fields (RFs) (Fig. 2). As a result, many responses were obtained with an RF either inside or outside the figure. For every pair of recording sites, three conditions could be discerned: both RFs of a pair of electrodes were within the boundaries of the figure ('in-in'), one RF was inside and one was outside ('in-out'), or both were outside on the background ('out-out') (Fig. 2). We could thus study synchrony for the same pairs of RFs either lying within, outside, or on either side of a texture

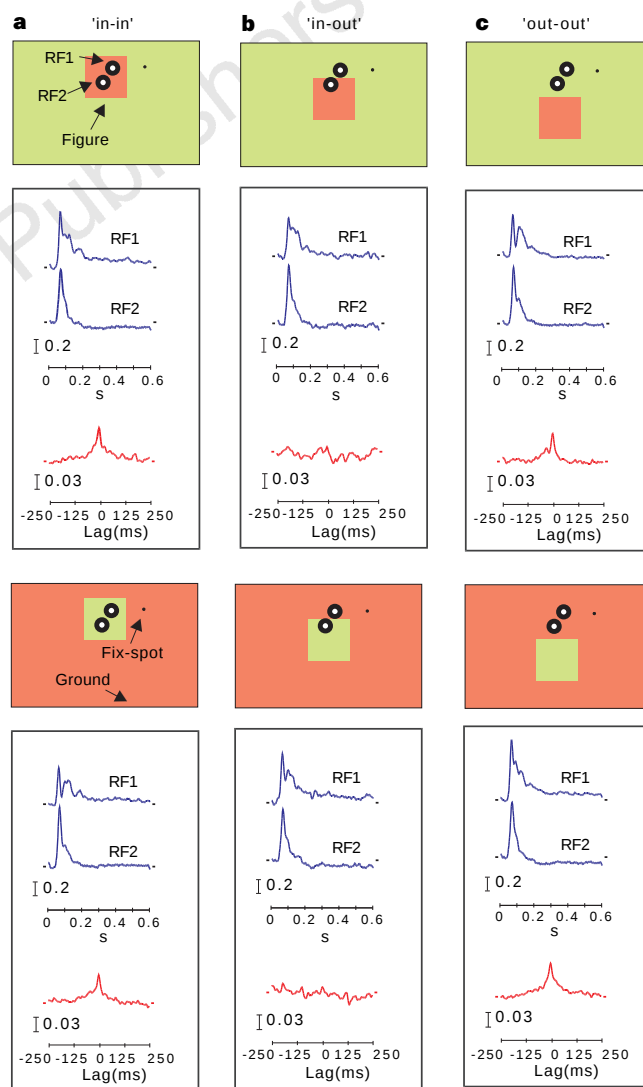


Figure 2 Complementary stimulus pairs, responses and cross-correlograms. Figure-ground stimuli were presented at various positions relative to the RFs (**a-c**). The figure could cover both RFs ('in-in'), one of the RFs, but not the other ('in-out'), or none of them, so that both were overlying the background ('out-out'). Different colours (orange or green) represent different local features, such as opposite directions of motion of random dots, or different orientations of line segments. Responses of (blue), and cross-correlations between (red), the multi-unit activity (MUA) of a pair of electrodes are shown for the 90° oriented texture stimulus. Response strength is given as fraction of peak response obtained at that electrode; correlation strength is given as normalized cross-correlation coefficient.

boundary. We used complementary stimulus pairs^{17–19}, where a particular orientation or direction of motion used for the figure in one trial was used for the background in another and vice versa (Fig. 2). In this way, responses were obtained for all possible combinations of local features present in the display. Data were analysed within 600 ms after stimulus onset, during which time the animals were fixating, and synchrony of firing between pairs of electrodes was assessed by calculating normalized cross-correlation functions.

If neuronal synchrony represents figure–ground segregation, one would expect synchrony (shown by a peak in the cross-correlogram) in the ‘in–in’ cases, where both RFs encode information about the same object. The ‘in–out’ cases, where the RFs encode different parts of the scene, would be expected to show less or no synchrony. It is not clear what should be expected from the ‘out–out’ case, where both RFs encode background features; these can be assumed to be of less ‘interest’ to the brain than figure features, which are important in, for example, object recognition.

Figure 2 shows the responses of, and cross-correlations between, two electrodes, for the oriented texture stimulus with a 90° difference between figure and ground. Both sites respond transiently to the onset of the stimulus, and almost equally well to the two orientations (compare the responses of each electrode in the two rows of panels). However, a flat correlogram is observed in both ‘in–out’ cases (Fig. 2b), whereas peaked correlograms are obtained for the ‘in–in’ and ‘out–out’ cases (Fig. 2a, c). The results obtained at this electrode pair are thus in good agreement with the above prediction.

Synchrony was generally found between many pairs of electrodes, even when their RFs were at a considerable distance (up to 4° of visual angle). Moreover, the given example shows that synchrony could be very stimulus-specific, rather than an artefact of eye movements for example. To analyse the population synchrony for all pairs of electrodes, and for all three stimulus display types, we took the correlation coefficient for each pair of electrodes at peak lag (peak lags of 95% of the correlograms ranged between –2.5 and +2.5 ms), and compared these for the three conditions ‘in–in’, ‘in–out’ and ‘out–out’. To lend equal weight to correlograms obtained with either orientation (or direction of motion), we averaged over both pairs of the complementary stimulus pairs. The results of this analysis are shown in Fig. 3. Points above the line where ‘in–in’ equals ‘in–out’ indicate stronger synchrony for the ‘in–in’ condition than for the ‘in–out’ condition, a result that would corroborate the synchrony hypothesis. Points below the line

would falsify the hypothesis. Figure 4a gives the population mean correlation coefficients for the three conditions and the three stimulus display types.

For the oriented texture with a 90° difference between figure and ground (Fig. 3a), many pairs show a result that is at least qualitatively similar to the result of the example pair (marked by an arrow). However, almost as many pairs show exactly the opposite result. The average peak-correlation coefficient (Fig. 4a) differs moderately between the three conditions. A significant difference is obtained between the conditions ‘in–out’ and ‘in–in’ ($P < 0.01$) and between ‘in–out’ and ‘out–out’ ($P < 0.001$, protected t - or Fisher LSD test, after one-way ANOVA $P = 0.00003$).

Figure 3b shows the results for the moving random dot stimulus with opposite directions of motion for figure and ground. Here, most pairs show equal synchrony for both conditions. Also, the mean correlation coefficients for all electrode pairs (Fig. 4a) are no longer significantly different between the ‘in–in’ and ‘in–out’ conditions (although ‘in–out’ versus ‘out–out’ is different, $P < 0.01$; after one-way ANOVA, $P = 0.02454$). A similar result is obtained for the oriented texture stimulus with a 20° difference between figure and ground (Fig. 3c). Most pairs show equal synchrony for the conditions ‘in–out’ versus ‘in–in’. Also, the three conditions result in the same mean synchrony (one-way ANOVA, $P = 0.68095$, NS) (Fig. 4a).

Although the complementary stimulus design ensures that all local features are used an equal number of times, it is important to note that global differences in figure–ground relationships (‘in–out’ versus ‘in–in’ or ‘out–out’) are accompanied by local differences in RF stimulation. More specifically, in the ‘in–out’ case there are always two opposite orientations or directions of motion stimulating either RF. Strictly, global stimulus properties cannot be changed without also changing local properties. Many previous studies on the role of synchrony have overcome this problem by ensuring that changes in local stimulation did not result in dramatic response differences^{11–16}. We have analysed response differences for the different sets of features, and excluded pairs from further analysis when responses to the two orientations or directions of motion differed by a factor of more than two in any of the two electrodes. These excluded pairs are shown as white circles in Fig. 3.

A second important issue is to what extent the stimuli match the RF tuning properties of the cells from which recordings were made. There might be differences between the synchrony of cells that are stimulated close to their preferred orientation and those which are less well stimulated²⁰. In addition, the random nature of the stimuli

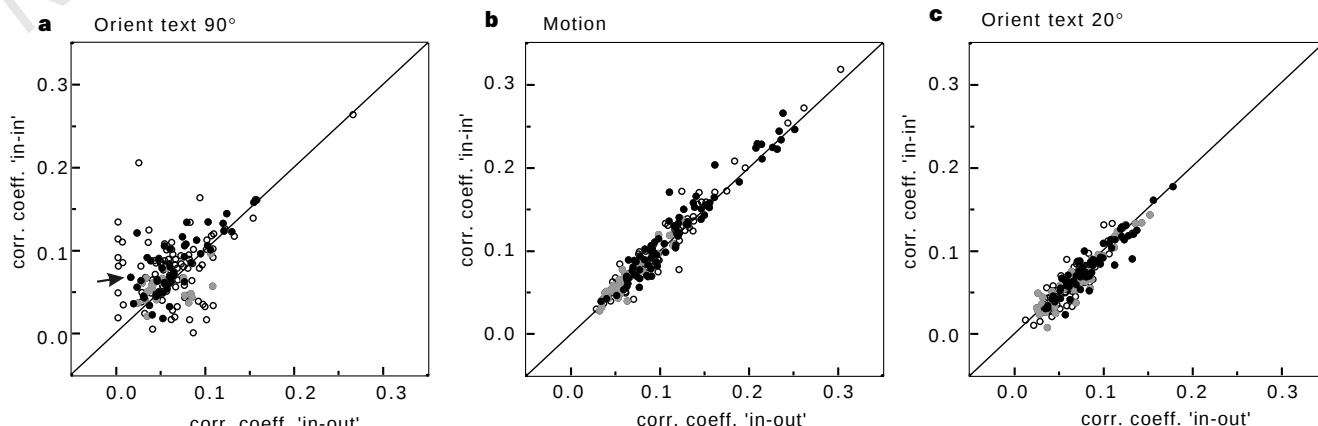


Figure 3 Population synchrony. **a–c**, All circles: normalized cross-correlation coefficients (NCCs) for the condition ‘in–out’ (abscissa) against normalized cross-correlation coefficients for the condition ‘in–in’ (ordinate), both at peak lag, for all pairs of recording sites that were sufficiently separate to allow for all three conditions to occur ($N = 172$). White circles: NCCs for those pairs for which the responses to the local features in any of the two electrodes differed by more than

a factor of 2. Grey circles: NCCs for pairs for which texture stimulation in any of the electrodes resulted in less than 40% of the response obtained for an optimally oriented bar. Black circles: NCCs for pairs remaining after the two previous classes were excluded (N values were 57, 78, 66 for the three stimulus types, respectively). The diagonal line shows where ‘in–in’ equals ‘in–out’. Arrowhead in **a** indicates the entry for the example pair of Fig. 2.

might cause cells not to be optimally stimulated, even when orientation of texture matches with the preferred orientation of the cells. Therefore, we also excluded from the analysis those pairs for which at any site the responses to the texture stimuli were less than 40% of the response to an optimally oriented bar (on average, responses to texture displays are ~59% of responses to optimal bars, but they are sometimes larger (14–189%; Fig. 3). Figure 3 shows these excluded pairs as grey and the finally remaining pairs as black circles.

Figure 4b shows the average synchrony for the three conditions in the remaining pairs. There are some subtle differences in overall synchronization strength, but the same (absences of) differences between the three conditions are observed as for those in Fig. 4a. Significant differences are only found between 'in-out', 'in-in' and 'out-out' cases for the 90° oriented texture stimulus.

Cells in V1 are not connected homogeneously: connections depend on relative distance and relative preferred orientation^{21–24}, which we also found (data not shown). To test whether only strongly interconnected cells show a correlate of figure-ground segregation in their synchrony, we isolated those pairs with a correlation strength for the 'out-out' cases that was greater than the median for that stimulus type (Fig. 4c). For the 90° oriented texture display,

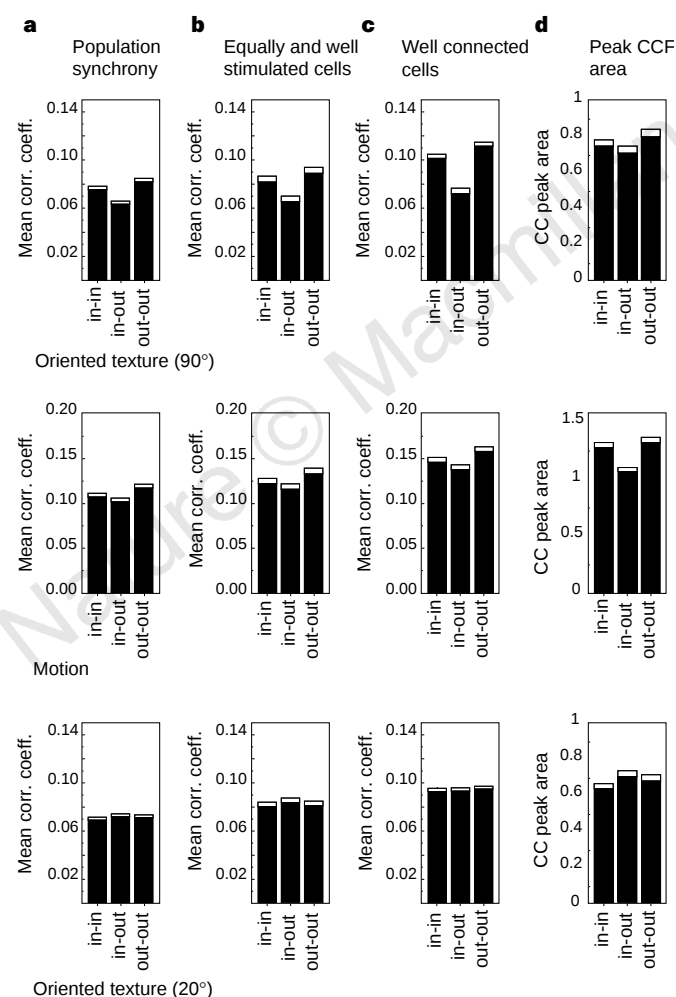


Figure 4 Mean population synchrony, and synchrony for specific populations of pairs. **a**, Average NCCs (at peak lag) for all pairs (all circles in Fig. 3) in the three conditions 'in-in', 'in-out' and 'out-out'. **b**, Mean NCCs for the set of pairs (black circles in Fig. 3) remaining after pairs with unequal stimulation for local features, or inadequate stimulation by texture stimuli, were excluded. **c**, Mean NCCs for pairs showing a correlation strength greater than the median for that stimulus. **d**, Mean 'area under the peak' of the correlograms for all sets of pairs. The white parts of the histograms represent the standard error of the mean. CCF, cross-correlation function.

this group of pairs shows a stronger difference between the condition 'in-out' and the conditions 'in-in' and 'out-out' than for the overall population (Fig. 4a). However, for both the motion stimulus and the oriented texture stimulus with a 20° difference between figure and ground, the results are not very different from the overall population results. The differences between 'in-in', 'in-out', and 'out-out' are not significant for the motion stimuli (except 'in-out' versus 'out-out', $P < 0.01$; after one-way ANOVA, $P = 0.02809$) or the 20° oriented texture stimuli (one-way ANOVA, $P = 0.85856$, NS).

In these analyses, we have taken the correlation coefficient at peak lag of the cross-correlogram as a measure of strength of synchrony between pairs. When taking zero lag instead, results are fully comparable (data not shown). Peak widths of the correlograms (width at half maximum height) differ for the three types of display. For the oriented texture displays (90° and 20°), median peak widths are 30 ms (range 0–100 ms). The distribution for the motion displays has a smaller range (20–80 ms) but shows more correlograms with large peak widths. We therefore also looked at a different measure of strength of synchrony, namely the area under the correlogram peak (Fig. 4d). In this measure, the only stimulus that results in a significant difference between the conditions 'in-in' and 'in-out' is the motion display ($P < 0.001$; after one-way ANOVA, $P = 0.0001$); both oriented texture displays show no significant correlate of figure-ground segregation (one-way ANOVA, $P = 0.26635$, NS and $P = 0.33152$, NS, respectively).

In conclusion, we observe clear neuronal synchrony between distant RFs in response to textured figure-ground stimuli. In general, however, the neuronal synchrony fails to reflect the perceptual organization of the scene: equal synchrony is observed between pairs of recording sites when both represent elements of the figure or when one represents the figure and the other the background. We were not able to segregate the results in any way that was related to response strengths, local receptive field tuning and stimulus matching, relative connection strength, or measure method of synchrony, to extract a distinct class of neuron pairs that was significantly influenced by the figure-ground configuration for all three displays.

A difference in synchrony is only observed between the three conditions when there is a 90° difference in orientation between the figure texture and the ground texture. This result could be interpreted as partial support for the hypothesis that synchrony is reflecting perceptual organization. A more parsimonious explanation, however, is to view synchrony as reflecting the horizontal connections within V1. It is known that these connections^{22,25,26} preferentially link sites that have similar orientation tuning^{21–23}. Therefore, less synchrony is to be expected when two sites are stimulated with orthogonal orientations, compared with when two sites are stimulated with identical ('in-in') or similar (20°) orientations. Opposite motions will also excite cells with similar orientation preferences, which will be equally well connected as cells responding to parallel motion.

Synchrony in V1 thus does not represent the binding of local features into segregating textures, but is instead a reflection of the horizontal connections in V1. By using texture segregation stimuli, we specifically investigated the role of synchrony in the grouping of local elements into figure and ground surfaces, in contrast to previous studies^{11–16} investigating the integration of contours. It remains to be seen whether these previous results can also be explained by local connectivity, or whether they still support a specific role for synchrony in the integration of contours. The general hypothesis that synchrony represents the binding of local features into global structures, however, is falsified by our results. □

Methods

Visual stimulation and behavioural control. Stimulus equipment and eye-movement recording system were as previously described^{27,28}. Trial initiation consisted of the appearance of a 0.2° red fixation spot. Monkeys were trained to maintain fixation on this spot. The stimulus, consisting of oriented line

segments or moving random dots that were organized so that a 4° square figure could be segregated from a background subtending 28° × 21° of visual angle, appeared on the screen 300 ms after fixation. One stimulation sequence occurred per trial. Animals were rewarded for maintaining fixation until the fixation spot turned off (600 ms after stimulus onset), and subsequently making a saccadic eye movement to the position of the square figure. All experimental procedures were approved by the institutional committee on animal welfare and were in accordance with NIH guidelines.

For the oriented line segment stimuli, stimulus onset consisted of the abrupt transition from a texture of randomly oriented line segments to a texture with either a 90° or a 20° orientation difference between figure and ground (shown in the top and bottom rows of Fig. 1). Line segments were 0.6° long and 0.03° wide; density was 11 line segments per square degree. The fraction of correct responses was the same for both orientation differences: 0.892 (±0.017) in both cases.

The motion stimuli consisted of a static random dot pattern (Fig. 1, centre row), where, in a 30-ms period, dots inside and outside the square-shaped region moved in opposite directions (over 0.06°). Half the pixels were black and half white; dot size was 0.03°. We have shown previously that this 'motion impulse' stimulus is effective in driving single units in a way that is comparable to the sudden onset of oriented texture stimuli, to which we wanted to compare the results. Responses are typically transient and can show direction selectivity¹⁹.

Recording of neuronal activity and data analysis. Multi-unit activity was recorded using electrodes implanted in area 17 (V1) in two macaque monkeys²⁷. Sixteen channels were recorded simultaneously in each animal. Eccentricity of the aggregate receptive fields of the neurons contributing to each channel ranged from 1.3° to 5.45°, and size from 0.18° to 1.4° (mean size, 0.52°). Median orientation selectivity ratio (optimally oriented bar response divided by least effective) was 1.94 (mean, 2.19; range, 1.16 to 9.03). All recording sites could be driven from either of the two eyes. There was no strong ocular dominance, as has been reported for layer 4C cells². Taking the RF sizes, tuning ratio and ocular dominance together, we conclude that the electrodes sampled neuronal activity over a distance of ~200–300 µm.

Synchrony of firing between all possible combinations of two electrodes with RFs sufficiently separate to allow for all three conditions 'in-in', 'in-out', and 'out-out' ($n = 172$) was assessed by computing the normalized cross-correlation coefficients²⁹. Cross-correlations were calculated over the full 600-ms time window following stimulus onset, during which time the monkeys maintained fixation, and after which they responded with a correct eye movement to the figure position. Average rate was subtracted from the responses before calculation. To remove stimulus-locked components from the cross-correlations, we used the technique of linear partialization of the cross-correlations²⁹.

Received 26 June; accepted 22 September 1998.

- Hubel, D. H. & Wiesel, T. N. Receptive fields and functional architecture of monkey striate cortex. *J. Physiol. (Lond.)* **195**, 215–243 (1968).
- Hubel, D. H. & Wiesel, T. N. Ferrier lecture: functional architecture of macaque monkey visual cortex. *Proc. R. Soc. Lond. B* **198**, 1–59 (1977).
- Schiller, P. H., Finlay, B. L. & Volman, S. F. Quantitative studies of single cell properties in monkey striate cortex I–V. *J. Neurophysiol.* **39**, 1288–1374 (1976).
- Maunsell, J. H. R. & Newsome, W. T. Visual processing in monkey extrastriate cortex. *Annu. Rev. Neurosci.* **10**, 363–401 (1987).
- Milner, P. M. A model for visual shape recognition. *Psychol. Rev.* **81**, 521–535 (1974).
- von der Malsburg, C. & Schneider, W. A neural cocktail-party processor. *Biol. Cybern.* **54**, 29–40 (1986).
- Eckhorn, R. et al. Coherent oscillations: a mechanism for feature linking in the visual cortex? *Biol. Cybern.* **60**, 121–130 (1988).
- Engel, A. K., König, P., Kreiter, A. K., Schillen, T. B. & Singer, W. Temporal coding in the visual cortex: new vistas on integration in the nervous system. *Trends Neurosci.* **15**, 218–226 (1992).
- Singer, W. Synchronization of cortical activity and its putative role in information processing and learning. *Annu. Rev. Physiol.* **55**, 349–374 (1993).
- Singer, W. & Gray, C. M. Visual feature integration and the temporal correlation hypothesis. *Annu. Rev. Neurosci.* **18**, 555–586 (1995).
- Gray, C. M., Engel, A. K., König, P. & Singer, W. Oscillatory responses in cat visual cortex exhibit intercolumnar synchronization which reflects global stimulus properties. *Nature* **338**, 334–337 (1989).
- Engel, A. K., König, P. & Singer, W. Direct physiological evidence for scene segmentation by temporal coding. *Proc. Natl Acad. Sci. USA* **88**, 9136–9140 (1991).
- Kreiter, A. K. & Singer, W. Stimulus dependent synchronization of neuronal responses in the visual cortex of awake macaque monkey. *J. Neurosci.* **16**, 2381–2396 (1996).
- Livingstone, M. S. Oscillatory firing and interneuronal correlations in squirrel monkey striate cortex. *J. Neurophysiol.* **75**, 2467–2485 (1996).
- Freiwald, W. A., Kreiter, A. K. & Singer, W. Stimulus dependent intercolumnar synchronization of single unit responses in cat area 17. *Neuroreport* **6**, 2348–2352 (1995).
- Brosch, M., Bauer, R. & Eckhorn, R. Stimulus dependent modulations of correlated high-frequency oscillations in cat visual cortex. *Cerebral Cortex* **7**, 70–76 (1997).
- Lamme, V. A. F., van Dijk, B. W. & Spekreijse, H. Texture segregation is processed by primary visual cortex in man and monkey. Evidence from VEP experiments. *Vision Res.* **32**, 797–807 (1992).

- Lamme, V. A. F., van Dijk, B. W. & Spekreijse, H. Contour from motion processing occurs in primary visual cortex. *Nature* **363**, 541–543 (1993).
- Lamme, V. A. F. The neurophysiology of figure-ground segregation in primary visual cortex. *J. Neurosci.* **15**, 1605–1615 (1995).
- König, P., Engel, A. K., Roelfsema, P. R. & Singer, W. How precise is neuronal synchronization? *Neur. Comput.* **7**, 469–485 (1995).
- Ts'o, D. Y., Gilbert, C. D. & Wiesel, T. N. Relationships between horizontal interactions and functional architecture in cat striate cortex as revealed by cross-correlation analysis. *J. Neurosci.* **6**, 1160–1170 (1986).
- Gilbert, C. D. & Wiesel, T. N. Columnar specificity of intrinsic horizontal and cortico-cortical connections in cat visual cortex. *J. Neurosci.* **9**, 2432–2442 (1989).
- Malach, R., Amir, Y., Harel, M. & Grinvald, A. Relationship between intrinsic connections and functional architecture revealed by optical imaging and *in vivo* targeted biocytin injections in primate striate cortex. *Proc. Natl Acad. Sci. USA* **90**, 10469–10473 (1993).
- Krüger, J. & Aiple, F. The connectivity underlying the orientation selectivity in the infragranular layers of monkey striate cortex. *Brain Res.* **477**, 57–65 (1989).
- Gilbert, C. D. & Wiesel, T. N. Clustered intrinsic connections in cat visual cortex. *J. Neurosci.* **3**, 1116–1133 (1983).
- Gilbert, C. D. Circuitry, architecture and functional dynamics of visual cortex. *Cerebral Cortex* **3**, 373–386 (1993).
- Lamme, V. A. F., Zipser, K. & Spekreijse, H. Figure-ground activity in primary visual cortex is suppressed by anesthesia. *Proc. Natl Acad. Sci. USA* **95**, 3263–3268 (1988).
- Bour, L. J., van Gisbergen, J. A. M., Bruijns, J. & Ottes, F. P. The double magnetic induction method for measuring eye movement—results in monkey and man. *IEEE Trans. Biomed. Eng.* **31**, 419–427 (1984).
- van der Togt, C., Lamme, V. A. F. & Spekreijse, H. Functional connectivity within the visual cortex of the rat shows state changes. *Eur. J. Neurosci.* **10**, 1490–1507 (1998).

Acknowledgements. We thank W. Singer, P. Roelfsema, K. Zipser and C. van der Togt for comments on an earlier version of the manuscript; K. Brandsma and J. de Feiter for biotechnical support; and P. Brassinga and H. Meester for technical support. This work was supported by a grant from the Royal Netherlands Academy of Arts and Sciences (KNAW) to V.A.F.L.

Correspondence and requests for materials should be addressed to V.A.F.L. (e-mail: v.lamme@amc.uva.nl).

Ethanol consumption and resistance are inversely related to neuropeptide Y levels

Todd E. Thiele*, Donald J. Marsh†, Linda Ste. Marie†, Ilene L. Bernstein* & Richard D. Palmiter†

* Department of Psychology, University of Washington, Box 351525, Seattle, Washington 98195, USA

† Howard Hughes Medical Institute and Department of Biochemistry, University of Washington, Box 357370, Seattle, Washington 98195, USA

Genetic linkage analysis of rats that were selectively bred for alcohol preference identified a chromosomal region that includes the neuropeptide Y (NPY) gene¹. Alcohol-preferring rats have lower levels of NPY in several brain regions compared with alcohol-non-preferring rats². We therefore studied alcohol consumption by mice that completely lack NPY as a result of targeted gene disruption³. Here we report that NPY-deficient mice show increased consumption, compared with wild-type mice, of solutions containing 6%, 10% and 20% (v/v) ethanol. NPY-deficient mice are also less sensitive to the sedative/hypnotic effects of ethanol, as shown by more rapid recovery from ethanol-induced sleep, even though plasma ethanol concentrations do not differ significantly from those of controls. In contrast, transgenic mice that overexpress a marked NPY gene in neurons that usually express it have a lower preference for ethanol and are more sensitive to the sedative/hypnotic effects of this drug than controls. These data are direct evidence that alcohol consumption and resistance are inversely related to NPY levels in the brain.

NPY is a 36-amino-acid neuromodulator that is distributed throughout the nervous system^{4,5}. Central infusion of NPY has shown that this peptide can promote feeding behaviour^{6,7}, reduce cerebrocortical excitability^{8,9} and anxiety-associated behaviours^{10,11}, and potentiate the sedative/hypnotic effects of certain drugs¹². We produced NPY-deficient (NPY^{−/−}) mice by gene targeting³. These

mice grow and reproduce at normal rates and show no abnormalities in their food intake or body weight, but they exhibit increased levels of anxiety and greater susceptibility to seizures than wild-type mice^{3,13}. Because NPY has been implicated in the high ethanol intake of rats selectively bred to model alcoholism (that is, the Indiana alcohol-preferring line^{1,14}), we studied the ethanol intake of NPY^{-/-} mice to determine whether increased ethanol consumption is a general characteristic of animals with reduced levels of or no NPY.

Mice were given continuous access to two bottles, one containing water and the other containing a 3% ethanol solution, for 8 days; the concentration of ethanol was then increased to 6%, 10% and 20% for further 8-day trials. The NPY^{-/-} mice consumed twice as much of the 6% ethanol solution (Fig. 1a, b) and 30–50% more of the two higher concentrations (Fig. 1b) than the wild-type controls. We expressed consumption of ethanol relative to total fluid consumption (ethanol-preference ratio). NPY^{-/-} mice showed a higher intake of ethanol and preferred ethanol to water (preference ratios greater than 0.50) during access to the 6% and 10% ethanol solutions

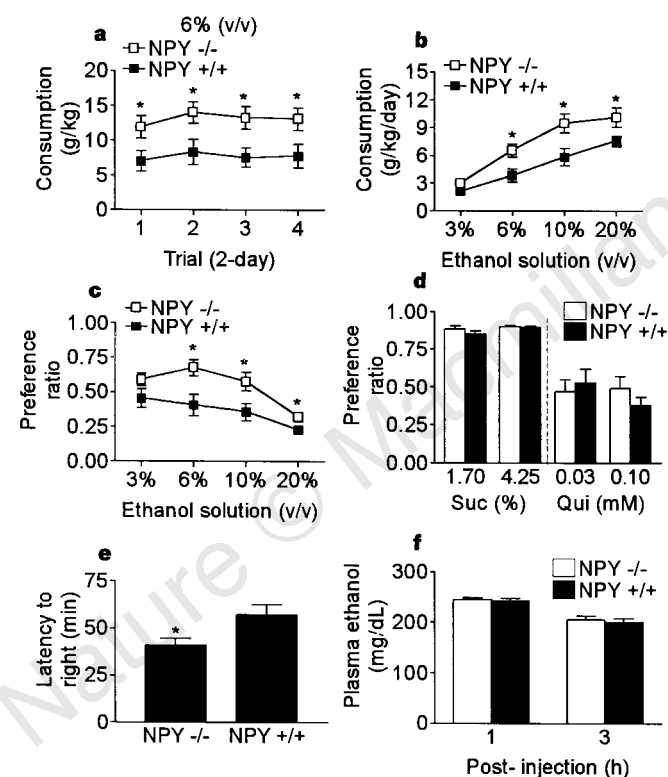


Figure 1 Consumption of ethanol and recovery from ethanol-induced sedation by NPY^{-/-} and wild-type (129/sv × C57BL/6; NPY^{+/+}) mice. **a**, Consumption (g kg⁻¹) of a 6% ethanol solution. **b**, Consumption (g kg⁻¹ d⁻¹) of each ethanol solution (8-day average). **c**, Ethanol-preference ratios (volume of ethanol consumed/total volume of fluid consumed) as a measure of relative ethanol preference. **d**, Preference ratios for sucrose (Suc) and quinine (Qui) (volume of taste solution consumed/total volume of fluid consumed). **e**, Time to regain the righting reflex (min), as a measurement of sensitivity to the sedative/hypnotic effects of ethanol (4.0 g kg⁻¹, i.p.). **f**, Plasma ethanol concentration (mg dl⁻¹) 1 h or 3 h after ethanol injection (4.0 g kg⁻¹, i.p.). All values are means ± s.e.m. ANOVAs indicate that the NPY^{-/-} mice drank significantly more ethanol ($P < 0.05$) and recovered from ethanol-induced sedation significantly sooner ($P < 0.05$) than NPY^{+/+} mice. NPY^{-/-} versus NPY^{+/+}, $P < 0.05$ (asterisks).

(Fig. 1c). There were no significant differences between genotypes nor were there differences over time in measures of average body weight (NPY^{-/-} mice, 32.78 ± 1.40 g; wild-type (NPY^{+/+}) mice, 33.11 ± 1.06 g), average food intake (NPY^{-/-} mice, 5.98 ± 0.18 g per day; NPY^{+/+} mice, 5.75 ± 0.17 g per day), or average total fluid consumption (ethanol solution plus water: NPY^{-/-} mice, 6.67 ± 0.33 ml per day; NPY^{+/+} mice, 6.39 ± 0.25 ml per day).

To determine whether the NPY^{-/-} mice exhibited general differences in taste preference, we tested separate groups of mice with sucrose and quinine solutions, using the same protocol as above. There were no significant differences between the genotypes in preference for these sweet and bitter compounds relative to water (Fig. 1d), nor were there differences in average total fluid consumption during access to sucrose (NPY^{-/-} mice, 6.20 ± 0.20 ml per day; NPY^{+/+} mice, 5.61 ± 0.49 ml per day) or quinine (NPY^{-/-} mice, 7.01 ± 0.52 ml per day; NPY^{+/+} mice, 7.63 ± 0.51 ml per day) solutions. Thus, increased preference for ethanol by NPY^{-/-} mice did not extend to other flavoured solutions. Furthermore, because ingestion

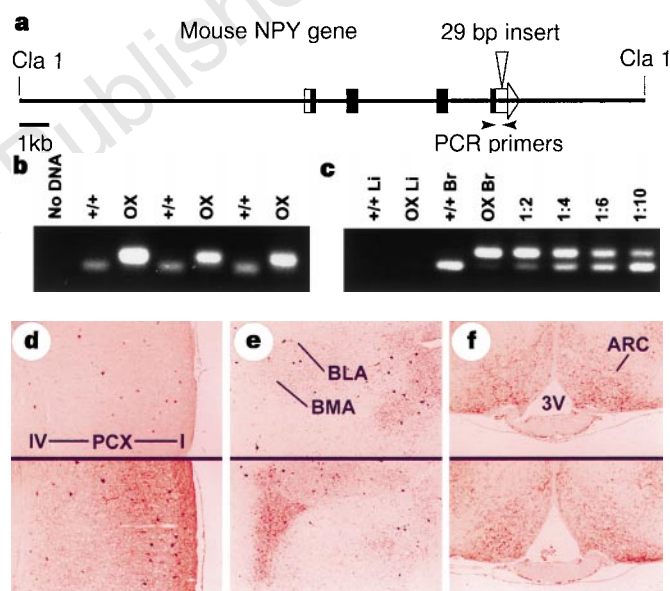


Figure 2 Generation and biochemical characterization of transgenic mice that overexpress a marked NPY gene. **a**, The NPY transgene construct. A DNA fragment containing ~10 kb 5' and ~5 kb 3' of the natural mouse NPY gene marked by a 29-bp oligonucleotide insertion into the 3'-untranslated region of the last exon was injected into fertilized FVB strain mouse eggs. **b**, PCR, using primers that flank the oligonucleotide insertion, was used to identify NPY transgenic mice and to estimate the transgene copy number. Amplification of the NPY transgene results in a larger, more abundant product than does amplification of wild-type NPY alleles. This is shown by comparison of several wild-type (+/+) and transgenic (OX) mice from one of the two high-copy-number lines that were established. **c**, Reverse transcription with PCR (RT-PCR), using primers that flank the oligonucleotide insertion, was used to estimate the level of liver and whole-brain NPY gene expression in wild-type and transgenic mice. Both wild-type and transgenic liver (Li) lack detectable NPY gene expression. NPY transgenic whole-brain (Br) complementary DNAs were either undiluted or diluted (1:2, 1:4, 1:6 and 1:10) with wild-type whole brain cDNAs, keeping the total cDNA per reaction constant. Amplification of transgenic NPY mRNAs results in a larger product than does amplification of wild-type NPY mRNAs. **d-f**, We used immunocytochemistry to compare the relative levels of NPY protein in different brain regions of wild-type (top panels) and transgenic (bottom panels) mice. NPY protein levels were higher, compared with wild-type mice, in regions of the cerebral cortex (**d**) and amygdala (**e**), but not the arcuate nucleus (**f**), of transgenic mice. PCX, parietal cortex, layers I to IV; BLA, basolateral amygdaloid nucleus; BMA, basomedial amygdaloid nucleus; ARC, arcuate nucleus; 3V, third ventricle. RT-PCR and immunohistochemistry procedures have been described previously³.

of food and sucrose solutions (which are both caloric) did not differ between the genotypes, increased intake of ethanol by NPY^{-/-} mice does not appear to be calorie-driven.

Because NPY may be involved in potentiating the sedative/hypnotic effects of drugs¹², we determined whether the sedative effects of ethanol¹⁵ are less in NPY^{-/-} mice compared with in wild-type mice. NPY^{-/-} mice were resistant to the sedative effects of ethanol, regaining their righting reflex ~15 min sooner than wild-type mice (Fig. 1e). Differences in sensitivity to the sedative/hypnotic effects of ethanol (and ethanol consumption) do not appear to be secondary to differences in acute clearance of ethanol, as plasma ethanol concentrations at 1 h and 3 h after ethanol administration did not differ between the genotypes (Fig. 1f).

To further study the relationship between NPY levels and ethanol, we examined transgenic mice that overexpress a marked NPY gene (NPY-OX) by using the same measurements as above. We generated transgenic NPY-OX mice by inserting a pair of oligonucleotides into the 3'-untranslated region of the last exon and then injecting a DNA fragment that contains ~10-kilobase (kb) 5' and ~5-kb 3' flanking regions of the mouse NPY locus (Fig. 2a) into fertilized mouse eggs (FVB strain). Gene dosage was estimated by the polymerase chain reaction (PCR) using oligonucleotides that flank the insert (Fig. 2b). We identified several founders with high copy number and established two lines that transmitted the transgene. Gene expression was estimated by PCR of reverse-transcribed brain RNA using primers flanking the insert (Fig. 2c). There was roughly five times more NPY transgene messenger RNA than endogenous NPY mRNA in the line of mice used here. Immunocytochemistry showed that, in the transgenic line, NPY protein was more abundant than normal in many regions of the brain that normally express NPY, including cortex (Fig. 2d), amygdala (Fig. 2e) and hippocampus (data not shown), but transgene expression was not apparent in the arcuate nucleus of the hypothalamus (Fig. 2f) or adrenal medulla (data not shown).

NPY-OX mice drank significantly less ethanol at all the concentrations tested (Fig. 3a, b) and had a lower relative preference for ethanol compared with wild-type littermates (Fig. 3c). NPY-OX mice were also more sensitive to ethanol-induced sedation than NPY^{+/+} mice (Fig. 3d) even though plasma ethanol concentrations

did not differ between genotypes 1 h and 3 h after ethanol administration (data not shown). Mice did not differ in average body weight (NPY-OX, 36.19 ± 0.90 g; NPY^{+/+} mice, 33.37 ± 0.86 g), average food intake (NPY-OX mice, 5.24 ± 0.16 g per day; NPY^{+/+} mice, 4.99 ± 0.08 g per day), or average total fluid consumption (ethanol solution plus water: NPY-OX mice, 7.08 ± 0.29 ml per day; NPY^{+/+} mice: 6.93 ± 0.21 ml per day).

The bidirectional relationship between NPY levels and ethanol intake is striking. A lack of NPY in knockout mice was associated with high ethanol consumption and low sensitivity to ethanol, whereas NPY overexpression in transgenic mice was associated with low ethanol consumption and high sensitivity to ethanol. Our results indicate that alcohol consumption and resistance are inversely related to NPY levels in the brain. Interestingly, the FVB strain wild-type mice appeared to consume more 20% ethanol solution (Fig. 3b) and required more time to recover from ethanol-induced sedation (Fig. 3d) than the wild-type mice (C57BL/6 × 129/Sv strain hybrids) in the knockout study (Fig. 1b, e). Inbred mouse strains differ markedly in their voluntary ethanol intake and ethanol resistance^{16,17}. There are many possible causes of these differences¹⁸, but it will be interesting to study NPY levels in mouse strains that differ greatly in alcohol consumption and/or alcohol resistance.

NPY is an inhibitory neuromodulator that acts widely in the brain through Y1, Y2, and Y5 receptors, all of which couple to heterotrimeric G proteins that inhibit production of cyclic AMP^{19,20}. Signalling through cAMP may mediate many responses to drugs and changes in this signalling pathway may mediate drug tolerance and dependence^{21–23}. One possible physiological role of NPY may be to inhibit cAMP production in response to alcohol, perhaps as a mechanism to limit alcohol intake. Thus, excess NPY would effectively reduce alcohol consumption whereas a lack of NPY would lead to increased consumption. One would also predict that neuromodulators that increase cAMP levels in critical neurons would also increase alcohol consumption. Consequently, removal (or antagonism) of such neuromodulators would result in decreased alcohol consumption and greater sensitivity to the sedative effects of ethanol. Inactivation of the *Drosophila amnesiac* gene, which encodes a secreted neuropeptide that stimulates cAMP production, renders flies more sensitive to ethanol-induced sedation²⁴. The protein encoded by *amnesiac* is related to mammalian growth-hormone-releasing factor and pituitary adenylcyclase-activating peptide. Thus, ethanol's effects in mammals and flies may be modified by counteracting actions of different neuromodulators that regulate cAMP in critical neuronal pathways.

NPY is not overexpressed in all regions of the brain. The lack of NPY transgene expression in the arcuate nucleus of the hypothalamus (Fig. 2f), a region thought to mediate food consumption^{13,25}, indicates that the effects on ethanol intake are probably not linked to calorie intake. We have considered the possibility that NPY

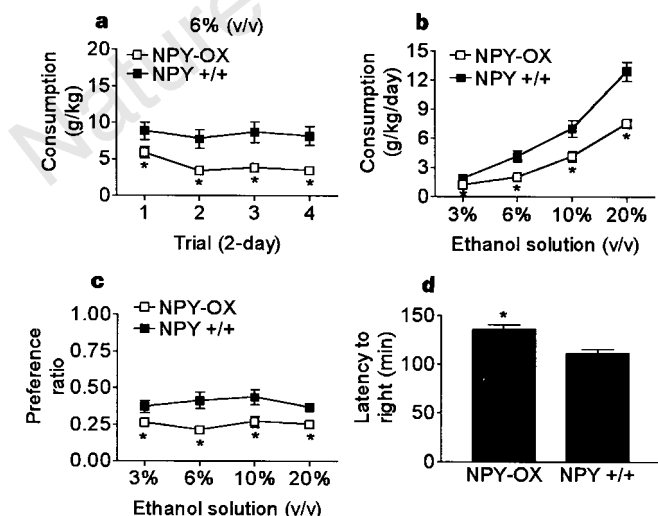


Figure 3 Ethanol consumption and recovery from ethanol-induced sedation by NPY-overexpressing mice (NPY-OX) and wild-type littermates (FVB; NPY^{+/+}). **a**, Consumption (g kg⁻¹) of a 6% ethanol solution. **b**, Consumption (g kg⁻¹ d⁻¹) of each ethanol solution (8-day average). **c**, Ethanol-preference ratios (volume of ethanol consumed/total volume of fluid consumed). **d**, Time to regain the righting reflex (min) following ethanol injection (4.0 g kg⁻¹; i.p.). All values are means ± s.e.m. ANOVAs show that the NPY-OX mice drank significantly less ethanol ($P < 0.001$) and required significantly more time to recover from ethanol-induced sedation ($P < 0.01$) than the NPY^{+/+} mice. NPY-OX versus NPY^{+/+}, $P < 0.05$ (asterisks).

Table 1 Excess NPY expression has no effect on anxiety

	FVB controls	NPY-OX
Elevated plus-maze test		
Number of mice	28	33
Open entries	10.14 ± 0.59	11.27 ± 0.61
Closed entries	12.07 ± 0.83	12.42 ± 0.74
Time in open (s)	101.6 ± 8.9	100.8 ± 7.6
Time in closed (s)	90.39 ± 5.4	95.42 ± 4.1
Contextual fear conditioning		
Number of mice	16	18
Time 'frozen' (%)	10.96 ± 2.6	14.43 ± 2.8

The elevated plus-maze test was done with 9- to 12-week-old male and female mice as described²⁹. Entries and time in the open and closed arms were measured for 5 min. Fear conditioning was tested as described³⁰ with some modifications. Briefly, on day 1, 12- to 15-week old male and female mice were placed in a Coulbourn Solid State Shocker and Model E1010 Modular Test Cage for 6 min. After becoming acclimatized to the chamber for 2 min, the mice received three footshocks (2 s, 0.65 mA; 1 min apart). On day 2, the mice were placed into the chamber and contextual fear conditioning was measured by scoring 'freezing behaviour' every 8 s for 6 min. Freezing is defined as the absence of all visible body and vibrissae movement except for that required for respiration. Freezing time is reported as a per cent of total observations. All values reported are mean ± s.e.m.

expression in the amygdala influences ethanol consumption by regulating anxiety. Central administration of NPY reduces anxiety and NPY^{-/-} mice score high on measures of anxiety^{10,11,13}. Because ethanol has anxiolytic properties²⁶, perhaps NPY-deficient mice consume more ethanol to modulate anxiety. However, the NPY-OX mice, which express more NPY in the amygdala, do not differ from their wild-type controls in anxiety measures such as fear conditioning and elevated plus maze tests (Table 1) yet drink less ethanol than the wild-type controls. Thus, differences in central NPY levels can be associated with differences in ethanol intake independently of any obvious differences in anxiety.

Two of the most obvious phenotypes described so far for NPY^{-/-} mice are increased sensitivity to seizures and increased ethanol consumption. The available data indicate that a major physiological function of NPY is to protect neural circuits from excessive stimulation. Thus, reduced levels of NPY would create a relatively unprotected neuronal system; this situation might be temporarily corrected by consumption of agents, such as ethanol, that can suppress neuronal activity in the cortex and hippocampus^{27,28}. It will be important to determine whether NPY expression is regulated by ethanol and where NPY acts to regulate ethanol consumption. □

Methods

Animals. The NPY-deficient mice were developed as described³. We used male NPY^{-/-} and wild-type mice that were maintained on a 129/sv and C57BL/6 hybrid strain (generations F₃–F₅ for each study). Transgenic mice were maintained on the FVB strain and males were used in all studies. We used naive mice in each study described below. Mice were individually housed in plastic mouse cages with free access to standard rodent chow (Teklad) and water throughout the experiments. The colony room was maintained at ~22 °C with a 12 h:12 h light:dark cycle.

Alcohol-intake test. Throughout the experiments, measures of fluid intake, food intake and body weight were taken every 2 days. NPY^{-/-} ($n = 11$) and wild-type ($n = 11$) mice were habituated in their home cage to drinking from two bottles containing water over 6 days. Mice were then given 24-h access to two bottles, one containing water and the other containing ethanol in water. The concentration of ethanol (v/v) was increased every 8 days; mice received 3%, 6%, 10% and finally 20% ethanol over the course of the experiment. The positions of the bottles were changed every 2 days to control for position preferences. These procedures were also used with NPY-OX mice ($n = 14$) and their wild-type littermates ($n = 14$). We calculated the average ethanol consumption per day for each ethanol concentration. To obtain a measure of ethanol consumption that corrected for individual differences in mouse size, we calculated grams of ethanol consumed per kg body weight per day for each mouse. As a measure of relative ethanol preference, we calculated ethanol-preference ratios at each ethanol concentration by dividing the volume of ethanol solution consumed by the total volume of fluid (ethanol plus water) consumed. Two-way, 2×4 (genotype $\times$ concentration) repeated-measure analyses of variance (ANOVAs) were used for statistical study of the data.

Sucrose- and quinine-preference test. NPY^{-/-} ($n = 7$) and wild-type ($n = 8$) mice were habituated in their home cage to drinking from two bottles containing water for 6 days. Over the next 8 days, mice were given plain water in one bottle and sucrose or quinine in the other bottle. The compounds were presented in the following order: sucrose solutions (1.70% and 4.25%) followed by quinine solutions (0.03 mM and 0.10 mM). Mice had 48-h access to each solution and the position of the solution was counterbalanced between animals. Preference for each solution was assessed by dividing the volume of taste solution consumed by the total volume of fluid (water plus taste solution) consumed to obtain a preference ratio. Data collected with each taste solution were analysed separately with two-way, 2×2 (genotype $\times$ concentration) repeated-measure ANOVA.

Test for sensitivity to the sedative/hypnotic effects of ethanol. NPY^{-/-} ($n = 7$) and wild-type ($n = 7$) mice were removed from their home cage and given an intraperitoneal (i.p.) injection of ethanol (4.0 g kg⁻¹, 20% (w/v) mixed in isotonic saline). At the onset of ethanol-induced sedation each mouse was placed on its back into a plastic U-shaped trough. The time (in min) that elapsed between the ethanol injection and when the mouse could right itself

onto all four paws, measured three times within a 30-s interval, was used as the index of time to regain the righting reflex. Using the same procedure, we studied NPY-OX mice ($n = 11$) and wild-type littermates ($n = 12$). These data were analysed with a one-way (genotype) ANOVA.

Plasma ethanol concentrations. Mice were given an i.p. injection of ethanol (4.0 g kg⁻¹, 20% (w/v) mixed in isotonic saline) and immediately returned to their home cage. One hour after ethanol injection, half of the NPY^{-/-} ($n = 7$) and half of the wild-type ($n = 7$) mice were rapidly anaesthetized with CO₂ and decapitated for blood collection. The remaining NPY^{-/-} ($n = 7$) and wild-type ($n = 7$) mice were anaesthetized and decapitated 3 h after the ethanol injection. Plasma ethanol concentrations were similarly determined in NPY-OX mice ($n = 7$ per time point) and wild-type littermates ($n = 7$ per time point). Plasma ethanol levels were determined by spectrophotometric methods (Sigma) and calculated as mg dl⁻¹. We used a two-way, 2×2 (genotype $\times$ time) repeated-measure ANOVA to analyse the data.

Received 28 July; accepted 18 September 1998.

1. Carr, L. G. *et al.* A quantitative trait locus for alcohol consumption in selectively bred rat lines. *Alcohol Clin. Exp. Res.* **22**, 884–887 (1998).
2. Ehlers, C. L. *et al.* Neuropeptide Y (NPY) levels in ethanol-naive alcohol preferring and non-preferring rats and in wistar following ethanol exposure. *Alcohol Clin. Exp. Res.* (in the press).
3. Erickson, J. C., Clegg, K. E. & Palmiter, R. D. Sensitivity to leptin and susceptibility to seizures of mice lacking neuropeptide Y. *Nature* **381**, 415–418 (1996).
4. Colmer, W. F. & Wahlestedt, C. *The Biology of Neuropeptide Y and Related Peptides* (Humana, Totowa, NJ, 1993).
5. Dumont, Y., Martel, J. C., Fournier, A., St Pierre, S. & Quirion, R. Neuropeptide Y and neuropeptide Y receptor subtypes in brain and peripheral tissues. *Prog. Neurobiol.* **38**, 125–167 (1992).
6. Clark, J. J., Karla, P. S., Crowle, W. R. & Karla, S. P. Neuropeptide Y and human pancreatic polypeptide stimulate feeding behavior in rats. *Endocrinology* **115**, 427–429 (1984).
7. Levine, A. S. & Morley, J. E. Neuropeptide Y: a potent inducer of consummatory behavior in rats. *Peptides* **5**, 1025–1029 (1984).
8. Woldbye, D. P., Madsen, T. M., Larsen, P. J., Mikkelsen, J. D. & Bolwig, T. G. Neuropeptide Y inhibits hippocampal seizures and wet dog shakes. *Brain Res.* **737**, 162–168 (1996).
9. Woldbye, D. P. *et al.* Powerful inhibition of kainic acid seizures by neuropeptide Y via Y5-like receptors. *Neurosci. Lett.* **3**, 761–764 (1997).
10. Hellig, M., McLeod, S., Koob, G. K. & Britton, K. T. Anxiolytic-like effect of neuropeptide Y (NPY), but not other peptides in an operant conflict test. *Regul. Pept.* **41**, 61–69 (1992).
11. Hellig, M. *et al.* Anxiolytic-like action of neuropeptide Y: mediation by Y1 receptors in amygdala, and dissociation from food intake effects. *Neuropsychopharmacol.* **8**, 357–363 (1993).
12. Yamada, K. *et al.* Neuropeptide Y reverses corticotropin-releasing hormone- and psychological stress-induced shortening of sodium pentobarbital-induced sleep in rats. *Brain Res.* **725**, 272–275 (1996).
13. Palmiter, R. D., Erickson, J. C., Holloper, G., Baraban, S. C. & Schwartz, M. W. Life without neuropeptide Y. *Recent Prog. Horm. Res.* **53**, 163–199 (1998).
14. Li, T.-K., Lumeng, L. & Doolittle, D. P. Selective breeding for alcohol preference and associated responses. *Behav. Genet.* **23**, 163–170 (1993).
15. Kurtz, D. L., Stewart, R. B., Zweifel, M., Li, T.-K. & Froehlich, J. C. Genetic differences in tolerance and sensitization to the sedative/hypnotic effects of alcohol. *Pharmacol. Biochem. Behav.* **53**, 585–591 (1996).
16. Crawley, J. N. *et al.* Behavioral phenotypes of inbred mouse strains: implications and recommendations for molecular studies. *Psychopharmacol.* **132**, 107–124 (1997).
17. Belknap, J. K., Crabbe, J. C. & Young, E. R. Voluntary consumption of ethanol in 15 inbred mouse strains. *Psychopharmacol.* **112**, 503–510 (1993).
18. Melo, J. A., Shendure, J., Pociask, K. & Silver, L. M. Identification of sex-specific quantitative trait loci controlling alcohol preference in C57BL/6 mice. *Nature Genet.* **13**, 137–138 (1996).
19. Hellig, M. & Widlerow, E. Neurobiology and clinical aspects of neuropeptide Y. *Crit. Rev. Neurobiol.* **9**, 115–136 (1995).
20. Gerald, C. *et al.* A receptor subtype involved in neuropeptide-Y-induced food intake. *Nature* **382**, 168–171 (1996).
21. Diamond, J. & Gordon, A. Cellular and molecular neuroscience of alcoholism. *Physiol. Rev.* **77**, 1–20 (1997).
22. Lane-Ladd, S. B. *et al.* CREB (cAMP response element-binding protein) in the locus coeruleus: biochemical, physiological, and behavioral evidence for a role in opiate dependence. *J. Neurosci.* **17**, 7890–7901 (1997).
23. Self, D. W. *et al.* Involvement of cAMP-dependent protein kinase in the nucleus accumbens in cocaine self-administration and relapse of cocaine-seeking behavior. *J. Neurosci.* **18**, 1848–1859 (1998).
24. Moore, M. S. *et al.* Ethanol intoxication in *Drosophila*: genetic and pharmacological evidence for regulation by the cAMP signaling pathway. *Cell* **93**, 997–1007 (1998).
25. Woods, S. C., Seeley, R. J., Porte, D. & Schwartz, M. W. Signals that regulate food intake and energy homeostasis. *Science* **280**, 1378–1383 (1998).
26. Stewart, R. B., Gatto, G. J., Lumeng, L., Li, T.-K. & Murphy, J. M. Comparison of alcohol-preferring (P) and nonpreferring (NP) rats on tests of anxiety and for the anxiolytic effects of alcohol. *Alcohol* **10**, 1–10 (1993).
27. Ryabinin, A. E., Criado, J. R., Henriksen, S. J., Bloom, F. E. & Wilson, M. C. Differential sensitivity of c-Fos expression in hippocampus and other brain regions to moderate and low doses of ethanol. *Mol. Psychiat.* **2**, 32–43 (1997).
28. Wright, J. M., Peoples, R. W. & Weight, F. F. Single-channel and whole-cell analysis of ethanol inhibition of NMDA-activated currents in cultured mouse cortical and hippocampal neurons. *Brain Res.* **738**, 249–256 (1996).
29. Rodgers, R. J. & Shepherd, J. K. Influence of prior maze experience on behaviour and response to diazepam in the elevated plus-maze and light/dark tests of anxiety in mice. *Psychopharmacol.* **113**, 237–242 (1993).
30. Kim, J. J. *et al.* Selective enhancement of emotional, but not motor, learning in monoamine oxidase A-deficient mice. *Neurobiology* **94**, 5929–5933 (1997).

Acknowledgements. We thank E. Sandgren for generating the NPY-OX mouse lines; G. Holloper and R. Cherne for help on maintaining and genotyping mice; I. Cubero and M. Miller for help in collecting data; and J. Seeley, L. Carr and T.-K. Li for comments and suggestions. This work was supported in part by the NIH and the University of Washington Alcohol and Drug Abuse Institute.

Correspondence and requests for materials should be addressed to T.E.T. (e-mail: thiele@u.washington.edu).

Control of neural crest cell fate by the Wnt signalling pathway

Richard I. Dorsky*, Randall T. Moon* & David W. Raible†

* Howard Hughes Medical Institute and Department of Pharmacology, and

† Department of Biological Structure, University of Washington School of Medicine, Seattle, Washington 98195, USA

Environmental signals are important in the development of neural crest, during which process multipotent progenitor must choose from several fates^{1–3}. However, the nature of these environmental signals is unknown. A previous fate map of zebrafish cranial neural crest showed that lineage-restricted clones of pigment cells arise from medial cells near the neural keel, and that clones of neurons arise from lateral cells farther from the neural keel⁴. *Wnt-1* and *Wnt-3a* are candidate genes for influencing neural crest fate, as they are expressed next to medial, but not lateral, crest cells. Here we determine the role of Wnt signals in modulating the fate of neural crest by injecting messenger RNAs into single, premigratory neural crest cells of zebrafish. Lineage analysis of injected cells shows that activation of Wnt signalling by injection of mRNA encoding cytoplasmic β -catenin promotes pigment-cell formation at the expense of neurons and glia. Conversely, inhibition of the Wnt pathway, by injection of mRNAs encoding either a truncated form of the transcription factor Tcf-3 or a dominant-negative Wnt, promotes neuronal fates at the expense of pigment cells. We conclude that endogen-

ous Wnt signalling normally promotes pigment-cell formation by medial crest cells and thereby contributes to the diversity of neural crest cell fates.

Wnt-1 and *Wnt-3a* are potential modulators of neural crest cell fate on the basis of their patterns of expression⁵ and their ability to alter cell fate during development⁶. Neural ectoderm is competent to express a neural crest marker, *slug*, before it is competent to express *Wnt-1* or *Wnt-3a*, indicating that these Wnts are probably not necessary for induction of neural crest⁷. The results of targeted disruption of these two Wnt genes in the mouse indicate that they may be essential in expansion of the neural crest⁸. However, ectopic Wnts increase the extent of neural crest markers in *Xenopus*^{9–11} when cell division is inhibited⁹, consistent with a function in addition to regulating neural crest cell proliferation. When we misexpressed *Wnt-1* in zebrafish by mRNA injection at the 1-cell stage, we saw a similar increase in amounts of neural crest markers and other morphological defects, such as deletion of forebrain structures (see Supplementary information). Although these embryos showed general increases in numbers of crest derivatives, such as pigment cells, such early global misexpression did not allow us to test whether gradients of Wnt signalling can directly affect the determination of neural crest cell fates.

As no previous study has specifically addressed whether Wnt signals promote specific neural crest cell fates at the expense of others, we investigated this role of Wnts in zebrafish neural crest. In the zebrafish, *Wnt-1* and *Wnt-3a* are expressed in the dorsal neural keel, from midbrain to spinal cord (Fig. 1a). Premigratory neural crest cells, visualized by a molecular marker, *forkhead-6* (ref. 12), are adjacent to this domain of Wnt expression (Fig. 1b). By labelling lateral cells that are destined to become neurons⁴ with fluorescein-

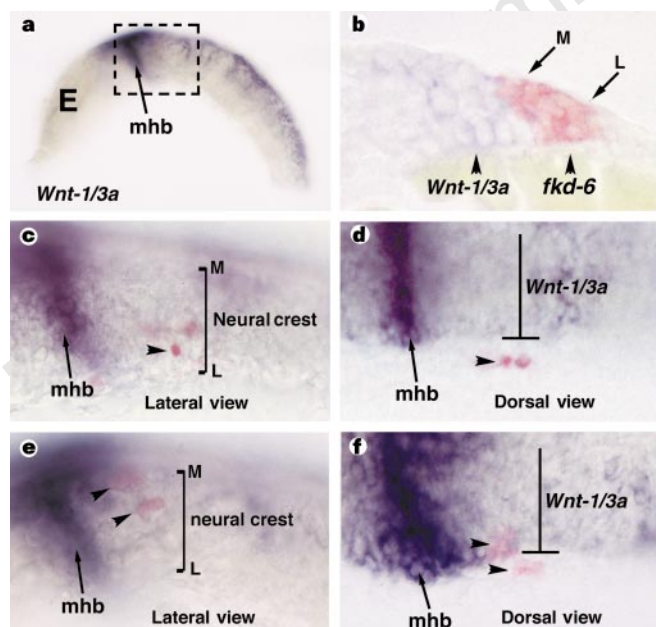


Figure 1 Analysis of the patterns of expression of *Wnt* genes in the vicinity of cranial neural crest. **a**, Lateral view of combined *Wnt-1*/*Wnt-3a* expression in a 12-h zebrafish embryo. Box indicates the area shown in **c–f**. E, eye; mhb, midbrain-hindbrain boundary. **b**, Cross-section through hindbrain region of a 12-hour-old zebrafish embryo double-stained for *Wnt-1*/*Wnt-3a* mRNA (blue) and *forkhead-6* mRNA (pink), a marker for premigratory neural crest¹². Medial neural crest cells (M), but not lateral neural crest cells (L), are adjacent to *Wnt*-expressing cells. **c**, **d**, Dye-labelled lateral neural crest cells (arrowhead), which produce neurons, are not next to *Wnt*-expressing cells in the neural keel. **e**, **f**, Dye-labelled medial neural crest cells (arrowheads), which produce pigment cells, contact *Wnt*-expressing cells. Compare **d**, **f**; note that in **d** the lateral cells (pink) are outside the domain of *Wnt* expression (denoted by inverted 'T'), whereas in **f** the medial cells (pink) overlap with or are adjacent to this domain. Anterior is to the left in **a**, **c–f**.

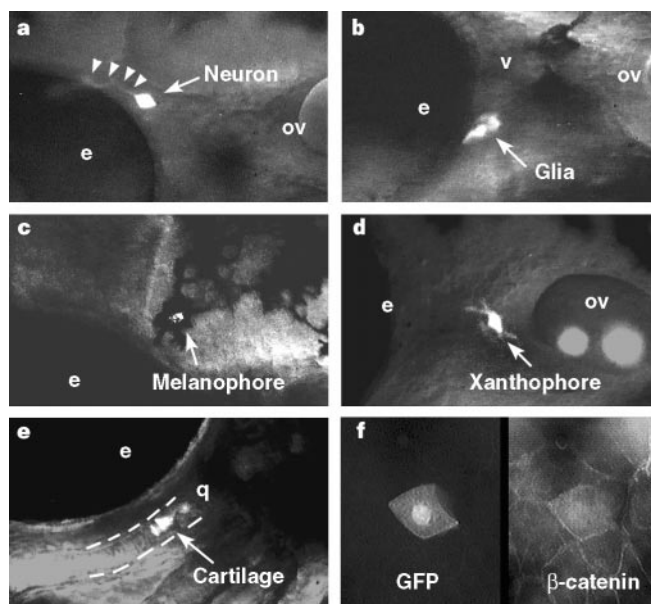


Figure 2 Individual zebrafish neural crest cells injected with GFP mRNA can be traced by GFP expression 24–48 h after injection. **a**, The neuron shown is large and round, is located in the trigeminal ganglion (v), and has an identifiable axon (arrowheads). e, eye; ov, otic vesicle. **b**, Clones of glia associated with peripheral nerves have small, oblong cell bodies which lie along axon pathways associated with ganglia (v). **c**, **d**, Pigment cells such as melanophores and xanthophores are identifiable by their pigment granules and stellate morphology. **e**, Cartilage cells have a characteristic cuboidal shape, with their long axis orthogonal to that of the cartilage rods (q, quadrate, outlined by dashed lines). **f**, A cell co-injected with mRNAs encoding GFP and β -catenin, and visualized by GFP fluorescence (left) and indirect immunofluorescence for β -catenin (right), shows increased amounts of cytoplasmic β -catenin compared with surrounding cells.

dextran, we found that they are several cell diameters away from the domain of Wnt expression (Fig. 1c, d). Dye-labelling of medial cells fated to become pigment⁴ showed that they are immediately adjacent to the *Wnt-1* and *Wnt-3a* expression domain (Fig. 1e, f). The correlation between cranial neural crest becoming pigment when nearest to Wnt-expressing cells, and becoming neurons when farther away from Wnt-expressing cells, led us to propose that Wnt signals may act to promote pigment-cell fate in the premigratory neural crest population.

We tested this hypothesis directly by overexpressing components of the Wnt pathway in individual neural crest cells. We micro-injected mRNA encoding green fluorescent protein (GFP) as a lineage tracer, mixed with mRNA encoding defined components of the Wnt pathway, into single premigratory lateral or medial neural crest cells using iontophoresis. GFP expression allows cells to be followed *in vivo*, and their fates to be classified by position and morphology. As seen in a previous fate map⁴, neural crest cells injected with GFP produced lineage-restricted clones (Fig. 2a–e). The progeny of GFP-injected cells adopted fates in similar proportions to those seen in the earlier fate map⁴ when followed for 24–48 hours after injection (Table 1 and Fig. 3a, b). Injection of mRNAs encoding GFP and an activated form of β -catenin, a cytoplasmic component of the Wnt signalling pathway⁶, followed by immunostaining for β -catenin, showed an increase in the amount of cytoplasmic β -catenin compared with uninjected cells (Fig. 2f). In contrast to GFP-injected controls, most lateral cells injected with activated β -catenin adopted pigment-cell fates and neurons or glia were produced rarely (Table 1 and Fig. 3a). Medial cells injected with β -catenin adopted pigment and cartilage fates at levels similar to GFP-injected controls (Table 1 and Fig. 3b), consistent with their location near an endogenous Wnt signal.

If Wnt signals are indeed involved in promoting medial cell fates at the expense of lateral fates, as our results indicate, then inhibition of Wnt signalling in medial cells should favour formation of neurons by neural crest. We therefore injected medial cells with mRNA encoding one of two inhibitors of Wnt signalling, either a mutant form of the transcription factor Tcf-3 (Δ Tcf) or a dominant-negative Wnt-1 (dnWnt-1). Both genes block endogenous Wnt signals and the effects of Wnt misexpression in *Xenopus*^{13–15} and zebrafish (data not shown). The effects of dnWnt-1, unlike those of Δ Tcf, can be rescued by overexpression of β -catenin, indicating that Wnt-1 and Tcf act at different points in the signalling cascade^{13,14}. In our experiments, these inhibitors had similar effects on neural crest cell fate (Table 1), showing the specificity of these reagents with regard to the Wnt pathway. Consistent with our hypothesis that Wnt signals produce medial fates in neural crest, injection of either inhibitor into medial neural crest resulted in a dramatic decrease in the production of pigment cells compared with control experiments (Table 1 and Fig. 3b). In addition, medial neural crest cells generated neurons when injected with the inhibitors, a fate never observed in control experiments or in previous fate-mapping studies⁴ (Table 1 and Fig. 3b). Injection of

the dominant-negative reagents had no effect on the fates of lateral cells, which are normally distant from Wnt signals in the central nervous system (Table 1 and Fig. 3a).

To test the specificity of injections of dnWnt-1, we attempted to rescue the effect of this construct by co-injecting β -catenin mRNA. The results obtained were identical to those produced with β -catenin alone, as injected cells mainly adopted pigment fates, regardless of the initial position of the injected cell (Table 1). Because the ability of dnWnt-1 to block pigment cell formation was completely inhibited by β -catenin, it is likely that dnWnt-1 interacts with an endogenous, β -catenin-dependent, Wnt pathway in neural crest cells. These data further support the hypothesis that *Wnt-1* and *Wnt-3a* expression in the neural tube promotes pigment-cell formation from neural crest.

Both our mRNA injections (Fig. 3a) and vital dye injections (data not shown) into lateral cells resulted in cartilage being labelled. Lateral cells never produced cartilage in the previous fate-mapping study⁴, but only superficial crest was labelled in these experiments so it is likely that deeper lateral cells account for the cartilage we observed. If cartilage arises from both medial and lateral cells, it is unlikely to be induced by localized *Wnt-1*/*Wnt-3a* signals. Consistent with this possibility, cartilage formation was not inhibited by either Δ Tcf or dnWnt-1 (Table 1 and Fig. 3a, b), indicating that Wnt signals are specifically required for promoting the pigment fate only. In contrast, few cells in any location produced glia when injected with β -catenin (Table 1 and Fig. 3a, b). Because glia arise from intermediate mediolateral positions⁴, they could be inhibited by the highest levels of Wnt signalling occurring directly adjacent to the

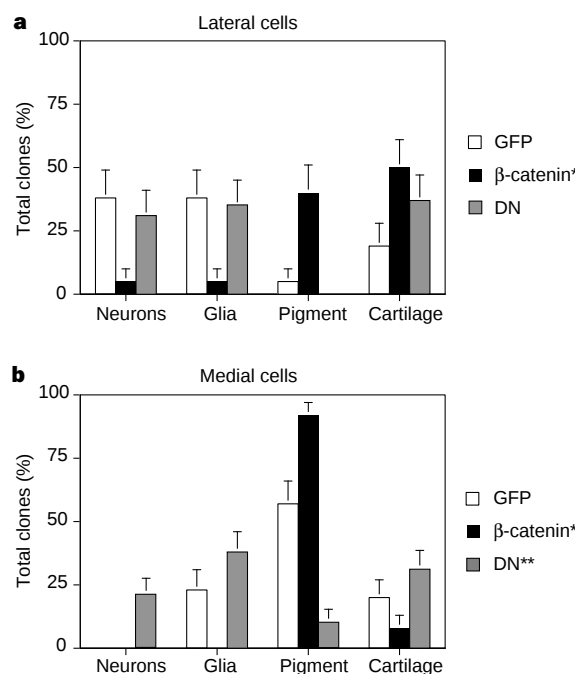


Figure 3 Manipulation of the Wnt pathway affects the fates of neural crest cells. **a**, Most lateral neural crest cells injected with β -catenin mRNA adopt pigment and cartilage fates, whereas controls expressing GFP alone mainly become neurons and glia. Lateral cells injected with dominant-negative reagents (DN, pooled data for Δ Tcf and dnWnt-1) behave like controls. GFP, $n = 21$ cells; β -catenin, $n = 20$ cells; DN, $n = 26$ cells. **b**, Medial neural crest cells injected with either Δ Tcf or dnWnt-1 (DN, pooled data) ectopically adopt neuronal fates at the expense of pigment formation. Cartilage formation is insensitive to Wnt inhibition with these reagents. Medial cells injected with β -catenin fail to adopt neuronal or glial fates and mainly generate pigment cells. GFP, $n = 30$ cells; β -catenin, $n = 26$ cells; DN, $n = 39$ cells. Error bars represent standard error: $\sqrt{(p(1-p)/n)}$; p is the proportion of cells in sample that adopt fate. Double asterisks indicate $P < 0.001$ vs GFP; single asterisks indicate $P < 0.01$ vs GFP; χ^2 analysis.

Table 1 Fates of injected neural crest cell clones

Position	Gene injected	Neurons	Glia	Pigment	Cartilage	Total	P
Lateral	GFP	8	8	1	4	21	–
	β -catenin	1	1	8	10	20	<0.001
	Δ Tcf	6	1	0	4	11	0.262
	dnWnt-1	2	8	0	5	15	0.282
	dnWnt + β -catenin	1	0	7	4	12	<0.001
Medial	GFP	0	7	17	6	30	–
	β -catenin	0	0	24	2	26	0.007
	Δ Tcf	3	5	1	8	17	0.002
	dnWnt-1	5	10	3	4	22	0.002
	dnWnt + β -catenin	0	0	13	1	14	0.048

P = P value compared to GFP-injected controls; calculated by χ^2 analysis.

Table 2 Mean clone sizes of injected cells $\pm$ s.e.m.

Gene injected	Neurons	Glia	Pigment	Cartilage
GFP	1.0 $\pm$ 0.0	2.7 $\pm$ 0.3	1.3 $\pm$ 0.1	3.0 $\pm$ 0.4
β -catenin	1.0*	1.0*	1.7 $\pm$ 0.1	3.8 $\pm$ 0.2
Δ Tcf	1.3 $\pm$ 0.2	2.0 $\pm$ 0.3	2.0*	3.2 $\pm$ 0.4
dnWnt-1	1.6 $\pm$ 0.2	3.3 $\pm$ 0.3	1.3 $\pm$ 0.3	4.3 $\pm$ 0.5
dnWnt + β -catenin	1.0*	—	1.5 $\pm$ 0.2	4.0 $\pm$ 1.1

Asterisk indicates no variance ($n = 1$).

neural keel; this situation is mimicked by β -catenin overexpression in our experiments.

Several observations lead us to conclude that our injections directly and specifically affected cell-fate determination. We did not see significant amounts of cell death following mRNA injections, ruling out the possibility that cells adopting particular fates were killed by gene misexpression. Activation of the Wnt pathway did not significantly increase the size of neural crest clones, and inhibition of Wnt signalling did not decrease proliferation (Table 2), indicating that if these genes are responsible for expanding the crest population they must do so at an earlier point in development and/or through a secondary signalling pathway. In addition, we did not observe any effects on cell migration, such as neurons located outside sensory ganglia, as the result of gene misexpression. We cannot exclude the possibility that other Wnt signals affect neural crest cells later in their development, but the specificity of our effects indicate that cell fate is likely to be determined early in this process, mediated by *Wnt-1* and *Wnt-3a*.

Our results support a model in which Wnt signalling promotes pigment-cell formation by cranial neural crest cells near the neural tube (Fig. 4). Neuronal and glial fates are inhibited by high levels of Wnt signalling in these medial cells, and can be promoted when Wnt signalling is blocked. Other signals must also be involved in promoting these alternative fates; candidates for a lateral signal include members of the bone morphogenetic protein family, which promote neuronal differentiation of neural crest *in vitro*^{16–18}. Our results showing that the cartilage fate is not medially restricted, and arises independently in the presence or absence of Wnt signals, support previous data that indicate that these precursors represent a separate lineage from other neural crest cells¹⁹.

Our data indicate that the fate of some neural crest cells is specified by local signals before they migrate away from the neural tube, and contrast with other models in which crest fate is determined by migratory environment or final destination²⁰. At the time of our injections, zebrafish cranial neural crest cells form lineage-restricted clones even after their fates are manipulated by modulation of Wnt signalling. Thus, despite individually generating only a single cell type, these cells are still multipotent.

In other vertebrates, and in the zebrafish trunk, there is a temporal

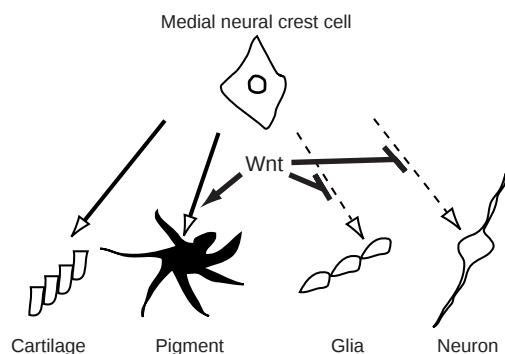


Figure 4 Model of Wnt function in specification of neural crest cell fate. Medial premigratory neural crest cells are adjacent to Wnt signals in the neural tube. These signals act upon the medial cells to promote pigment-cell fates at the expense of glial and neuronal fates. Cartilage production does not require Wnt signalling.

rather than a spatial difference in neural crest fates, with lineage-restricted clones becoming more prevalent over time^{3,21}. A detailed analysis of the relationship between time of migration and Wnt expression will help to determine whether this is a universal mechanism for neural crest patterning. This possibility is supported by the phenotypes of mice lacking *Wnt-1* and *Wnt-3a*; in these mice, pigment cells are the most severely depleted neural crest derivative⁸. A potential parallel role for Wnt proteins in invertebrates has been shown: *DWnt-2* regulates pigment-cell formation in the *Drosophila* testis²². It is possible that the molecular control of pigment-cell determination has been conserved even in species lacking neural crest. □

Methods

In situ hybridization and dye labelling. Digoxigenin- and fluorescein-labelled RNA probes were prepared from templates encoding *forkhead-6* (a gift from J. Odenthal¹²), *Wnt-1* and *Wnt-3a*. The *Wnt-3a* complementary DNA was produced by A. Lekven by 5' rapid amplification of cloned ends (RACE) (Clontech) with primers obtained from a partial cDNA provided by V. Korzh²³. Single and double whole-mount *in situ* hybridizations were performed as described²⁴, using BCIP/NBT and Fast Red (Boehringer) as substrates. Single-cell dye labelling²⁵ was detected using Fast Red.

mRNA injections. We produced mRNA using the Message Machine kit (Ambion) from templates containing *GFP*, *ptβ-catenin*²⁶, *ΔTcf*³ and *dnWnt-1* cDNAs in the vector pCS2+. Uncapped fluorescently labelled RNA was produced similarly, using a fluorescein-UTP mix (Boehringer), and was used to monitor mRNA iontophoresis. RNAs were mixed to a concentration of 0.25–1 mg ml⁻¹ each and combined with KCl at a final concentration of 0.1 M. We injected the RNA mixture iontophoretically into premigratory neural crest as described for fluorescently labelled dextrans⁴. At 1–2 h post-injection we examined embryos for GFP-expressing neural crest cells, and noted the position of these cells. Live cells were followed for 24–48 h, using a Videoscope ICCD camera, until their final position and morphology could be used to determine their fate⁴. We detected β -catenin expression using whole-mount immunostaining with polyclonal antiserum²⁶.

Received 21 May; accepted 21 September 1998.

- Baker, C. V. H., Bronner-Fraser, M., LeDourain, N. M. & Teillet, M. A. Early- and late-migrating cranial neural crest cell populations have equivalent developmental potential *in vivo*. *Development* **124**, 3077–3087 (1997).
- Shah, N. M., Marchionni, M. A., Isaacs, I., Stroobant, P. & Anderson, D. J. Glial growth factor restricts mammalian neural crest stem cells to a glial fate. *Cell* **77**, 349–360 (1994).
- Raible, D. W. & Eisen, J. S. Regulatory interactions in zebrafish neural crest. *Development* **122**, 501–507 (1996).
- Schilling, T. F. & Kimmel, C. B. Segment and cell type lineage restrictions during pharyngeal arch development in the zebrafish embryo. *Development* **120**, 483–494 (1994).
- Hollyday, M., McMahon, J. A. & McMahon, A. P. Wnt expression patterns in chick embryo nervous system. *Mech. Dev.* **52**, 9–25 (1995).
- Cadigan, K. M. & Nusse, R. Wnt signaling: a common theme in animal development. *Genes Dev.* **11**, 3286–3305 (1997).
- Dickinson, M. E., Selleck, M. A., McMahon, A. P. & Bronner-Fraser, M. Dorsalization of the neural tube by the non-neural ectoderm. *Development* **121**, 2099–2106 (1995).
- Ikeya, M., Lee, S. M. K., Johnson, J. E., McMahon, A. P. & Takada, S. Wnt signalling required for expansion of neural crest and CNS progenitors. *Nature* **389**, 966–970 (1997).
- Saint-Jeannet, J. P., He, X., Varmus, H. E. & Dawid, I. B. Regulation of dorsal fate in the neuraxis by *Wnt-1* and *Wnt-3a*. *Proc. Natl Acad. Sci. USA* **94**, 13713–13718 (1997).
- Chang, C. B. & Hemmati-Brivanlou, A. Neural crest induction by *Xwnt7B* in *Xenopus*. *Dev. Biol.* **194**, 129–134 (1998).
- LaBonne, C. & Bronner-Fraser, M. Neural crest induction in *Xenopus*: evidence for a two-signal model. *Development* **125**, 2403–2414 (1998).
- Odenthal, J. & Nüsslein-Volhard, C. *forkhead* domain genes in zebrafish. *Dev. Gene Evol.* **208**, 245–248 (1998).
- Molenaar, M. et al. *XTcf-3* transcription factor mediates beta-catenin-induced axis formation in *Xenopus* embryos. *Cell* **86**, 391–399 (1996).
- Hoppler, S., Brown, J. D. & Moon, R. T. Expression of a dominant-negative Wnt blocks induction of *MyoD* in *Xenopus* embryos. *Genes Dev.* **10**, 2805–2817 (1996).
- McGrew, L. L., Hoppler, S. & Moon, R. T. Wnt and FGF pathways cooperatively pattern anteroposterior neural ectoderm in *Xenopus*. *Mech. Dev.* **69**, 105–114 (1997).
- Reissmann, E. et al. Involvement of bone morphogenetic protein-4 and bone morphogenetic protein-7 in the differentiation of the adrenergic phenotype in developing sympathetic neurons. *Development* **122**, 2079–2088 (1996).
- Shah, N. M., Groves, A. K. & Anderson, D. J. Alternative neural crest cell fates are instructively promoted by TGFβ superfamily members. *Cell* **85**, 331–343 (1996).
- Varley, J. E. & Maxwell, G. D. BMP-2 and BMP-4, but not BMP-6, increase the number of adrenergic cells which develop in quail trunk neural crest cultures. *Exp. Neurol.* **140**, 84–94 (1996).
- LeDourain, N. M. & Teillet, M. A. Experimental analysis of the migration and differentiation of neuroblasts of the autonomic nervous system and of neuroectodermal mesenchymal derivatives, using a biological cell marking technique. *Dev. Biol.* **41**, 162–184 (1974).
- Bronner Fraser, M. & Fraser, S. E. Cell lineage analysis of the avian neural crest. *Development Suppl.* **2**, 17–22 (1991).

21. Henion, P. D. & Weston, J. A. Timing and pattern of cell fate restrictions in the neural crest lineage. *Development* **124**, 4351–4359 (1997).
22. Kozopas, K. M., Harryman Samos, C. & Nusse, R. *DWnt-2*, a *Drosophila* Wnt gene required for the development of the male reproductive tract, specifies a sexually dimorphic cell fate. *Genes Dev.* **12**, 1155–1165 (1998).
23. Krauss, S., Korzh, V., Fjose, A. & Johansen, T. Expression of four zebrafish wnt-related genes during embryogenesis. *Development* **116**, 249–259 (1992).
24. Jowett, T. & Lettice, L. Whole-mount *in situ* hybridizations on zebrafish embryos using a mixture of digoxigenin- and fluorescein-labelled probes. *Trends Genet.* **10**, 73–74 (1994).
25. Appel, B. *et al.* Motoneuron fate specification revealed by patterned LIM homeobox gene expression in embryonic zebrafish. *Development* **121**, 4117–4125 (1995).
26. Yost, C. *et al.* The axis-inducing activity, stability, and subcellular distribution of beta-catenin is regulated in *Xenopus* embryos by glycogen synthase kinase 3. *Genes Dev.* **10**, 1443–1454 (1996).

Supplementary information is available on Nature's World-Wide Web site (<http://www.nature.com>) or as paper copy from the London editorial office of Nature.

Acknowledgements. We thank G. Kruse for technical support; J. Yang-Snyder for help with related confirmatory experiments (not shown); J. Odenthal for *forkhead-6* cDNA; V. Korzh for the *Wnt-3a* cDNA; and A. Lekven for the *Wnt-3a* RACE clone. R.L.D. was supported by the Molecular Neurobiology Training Grant from the NIH and R.T.M. acknowledges support as an Investigator of the HHMI. This research was funded by the HHMI (R.T.M.) and the NIH (D.W.R.).

Correspondence and requests for materials should be directed to R.L.D. (e-mail: rdorsky@u.washington.edu).

Proto-oncogene *PML* controls genes devoted to MHC class I antigen presentation

Pan Zheng*, Yong Guo, Qingtian Niu, David E. Levy, Jacqueline A. Dyck†, Shengli Lu, Lori A. Sheiman & Yang Liu*

Department of Pathology and Kaplan Comprehensive Cancer Center, New York University Medical Center, New York, New York 10016, USA

† Department of Medicine, University of California at San Diego, La Jolla, California 92093, USA

Fragments of foreign antigens associated with class I molecules of the major histocompatibility complex (MHC) are presented at the cell surface to elicit an immune response. This presentation requires the coordinated expression of several genes contained in the MHC^{1–5}, including those encoding the MHC class I heavy chain, the proteins LMP-2 and LMP-7, which are involved in the proteasomal degradation of cytosolic antigens into peptide fragments that are destined for association with MHC class I molecules, and TAP-1 and TAP-2, which transport these fragments across the membrane of the endoplasmic reticulum at the start of their journey to the cell surface. In many virus-transformed cell lines^{6,7} and spontaneous tumours^{8–10}, these genes are simultaneously repressed. However, the key factor(s) that are essential for their expression and repression have not been identified. Here we report that the proto-oncogene product PML induces expression of LMP-2, LMP-7, TAP-1 and TAP-2 in an MHC-class I-negative, recurrent tumour, leading to the re-expression of cell-surface MHC in tumours and to rejection of the tumours. PML also regulates MHC expression in untransformed fibroblasts. We conclude that malfunction of PML may enable a tumour to evade the immune defence of its host.

We have previously shown that the plasmacytoma J558, when transfected with the co-stimulatory molecule B7-1, is rejected in immune-competent mice, and that the tumour does not recur in the overwhelming majority of mice that had rejected the tumour¹¹. However, tumours occasionally recur in a few mice after tumour rejection. Mice that have rejected the J558-B7 tumour have an expanded population of cytotoxic T-lymphocyte (CTL) precursors specific for the unmutated tumour antigen P1A, as do those with recurrent tumours (data not shown). Thus, the recurrent tumours

have arisen in the presence of strong anti-tumour immunity. The parental (J558-B7) and the recurrent (ReB7) tumours both express a comparable amount of P1A messenger RNA¹² (Fig. 1a). We could detect no expression by ReB7 of any cell-surface MHC class I loci, H-2D^d, L^d and K^d (Fig. 1b). As expected, ReB7 tumours are completely resistant to cytolysis by a CTL clone¹² specific for the tumour antigen P1A (Fig. 1c). Addition of exogenous P1A peptide restored recognition by CTLs (Fig. 1c).

This loss of MHC expression is probably due to a defect in antigen presentation. To test this idea, we analysed the expression of the MHC proteins TAP-1, TAP-2, LMP-2, LMP-7 and class I heavy chain by northern blotting. The mRNAs for TAP-1, TAP-2, LMP-2 and LMP-7 were barely detectable in ReB7 (Fig. 1d), consistent with the observed defects in antigen processing. Although there was some reduction in the amount of MHC class I heavy chain K^d mRNA, plenty was still detectable, consistent with the idea that the exogenous peptides can restore cell-surface MHC. As a control, we used RAW8.1, a B-cell leukaemia of BALB/c origin, which expresses the highest amount of all components devoted to MHC class I antigen processing. The recurrent tumour ReB7 thus evades immune recognition by downregulating the many molecules that are critical for MHC class I antigen processing.

We designed a strategy to identify genes that could compensate for the antigen-presentation defects in ReB7 cells (see Methods). After isolating a complementary DNA clone (F12) that reproducibly complemented the expression of H-2L^d in a transient fusion assay (Fig. 2a, left), we subcloned it into EBO-Sfi, a vector known to give high expression in plasmacytoma¹³, and used the vector to produce stable F12 transfectants. Clones that survived hygromycin selection were tested for cell-surface expression of MHC class I. In two independent transfections, 16 of the 25 EBO-Sfi-F12-transfected ReB7 clones expressed significant amounts of cell-surface MHC class I, as judged by flow cytometry, whereas none of the 13 clones transfected with vector alone expressed any (see Fig. 2a, right, for a representative clone in each group). mRNAs encoding TAP-1, TAP-2, LMP-2 and LMP-7 were also induced by expression of PML (Fig. 2b, right). In addition, F12-transfected ReB7 (clone 10A7) was killed by P1CTLs in the absence or presence of exogenous P1A peptide, whereas the ReB7 subline SG9, which was transfected by vector alone, was killed only in the presence of an exogenous P1A peptide (Fig. 2c). Moreover, most mice rejected the F12-transfected ReB7 cell line 10A7, and the kinetics of rejection were similar to those of parental J558-B7. In contrast, neither ReB7 nor the vector-transfected ReB7 subline SG9 was rejected (Fig. 2d). These results show that overexpression of F12 cDNA rendered the ReB7 tumour susceptible to the host immune response.

DNA sequencing revealed that F12 cDNA is an isoform of PML, a proto-oncogene with no known function. F12 cDNA comprises exons 1–4, 6 and 7 and part of exon 9 of the PML gene (Fig. 3a); this cDNA encodes a 65K protein, as shown by *in vitro* translation (Fig. 3b). The predicted amino-acid sequence encoded by PML-F12 is shown in Fig. 3c. PML-F12 protein has three putative zinc-finger boxes, four coiled-coil clusters, a putative activation domain, and a nuclear-localization sequence. The sequence of exons 1–4, 6 and 7 of F12 is essentially identical to that of the murine PML gene¹⁴; the absence of exon 8 causes a frameshift and early termination in exon 9, although the significance of this is not clear. The PML-F12 isoform was overexpressed in the stable PML-F12 transfectant 10A7 in comparison with SG9, which was transfected with vector alone (Fig. 2b, left). A PML-F12 probe detected 3-fold more PML mRNA in 10A7, whereas a PML exon-8 probe detected comparable amounts of PML in 10A7 and SG9.

Extensive analysis by northern blotting and polymerase chain reaction combined with reverse transcription (RT-PCR) indicated that the amount as well as the isoform of PML mRNA in ReB7 are essentially identical to those in J558-B7 (data not shown). In particular, comparable levels of PML-F12 mRNA are present in

* Present address: Department of Pathology, Ohio State University Medical Center, Columbus, Ohio 43210, USA.

both cell lines (data not shown), so an abnormality in PML gene expression is unlikely to cause malfunction of PML in ReB7. To determine whether an endogenous mutation could be responsible for the antigen-processing defects in ReB7, we amplified each of the nine PML exons by PCR from genomic DNA isolated from J558-B7 and from ReB7, and sequenced the PCR products. We also produced multiple, overlapping cDNA clones that covered all the PML coding regions. Our results indicate that exon 3 of PML has a lesion unique to ReB7, whereas the sequences in all other exons are identical in ReB7 and its parental cell lines (data not shown). To confirm the presence of a mutation in exon 3, we isolated several clones of the exon-3 DNA fragment from J558-B7 and a subline of ReB7. Two exon-3 clones from the ReB7 subline have an identical deletion (G at position 962 of the murine PML sequence¹⁴) (Fig. 3d, top panel). This deletion is not present in the other four exon-3 clones from the same ReB7 subline, indicating that one copy of the PML gene in ReB7 cells contains the deletion but the other is unchanged. None of the five exon-3 clones from J558-B7 (three clones from genomic DNA and two clones from cDNA) contains this mutation. The deletion at G962 causes a frameshift and translational termination of PML protein in exon 3 (Fig. 3d, lower panel). *In vitro* translation confirms that the mutant encodes a truncated protein (Fig. 3b).

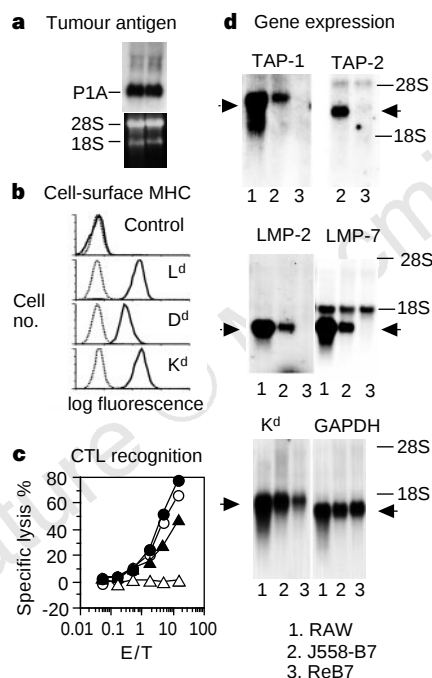


Figure 1 Multiple antigen-presentation defects in a recurring tumour cell line, ReB7. **a**, Normal expression of the tumour antigen P1A in ReB7. Total RNA isolated from J558-B7 and ReB7 was probed with P1A cDNA (top); RNA loading is calibrated against ribosomal RNA at the bottom. **b**, Expression of H-2L^d, D^d and K^d in ReB7 (dotted lines) and parental J558-B7 (solid lines) cells as determined by flow cytometry using specific monoclonal antibodies: HB27 (against L^d), HB76 (D^d), or HB159 (K^d). **c**, Cytotoxicity of ReB7 (triangles) or J558-B7 (circles) by the P1A-specific CTL line P1CTL (ref. 12) in the presence (black symbols) or absence (white symbols) of the exogenous peptide P1A (sequence, LPYLGLWLVF; 1 μ g ml⁻¹). **d**, Expression of LMP-2, LMP-7, TAP-1, TAP-2 and H-2K^d heavy chain genes (arrowheads); expression of the glyceraldehyde 3-phosphate dehydrogenase (GAPDH) gene was used as a loading control. E/T, effector-to-target ratio.

As we obtained multiple wild-type PML cDNA clones from a ReB7 subclone (data not shown), the wild-type alleles must have been expressed. Thus, the PML mutant may act as a dominant-negative regulator. To test this, we transfected PML-F12dG mutant cDNA into parental J558-B7 cells. One month after drug selection, most drug-resistant, PML-F12dG-transfected cells had significantly less cell-surface MHC, and ~15% of them had none at all (Fig. 3e, left). A large collection of stable clones obtained from the bulk culture expressed barely any cell-surface MHC (see Fig. 3e, right, for an example) and markedly reduced amounts of TAP-1, TAP-2, LMP-2 and LMP-7 mRNAs in comparison with J558-B7 transfected with vector alone (Fig. 3f). Thus, the frameshift mutation causes downregulation of MHC on the cell surface and acts dominantly. The PML-F12dG mutant contains the 5' proline-rich domain, three zinc-finger boxes, and most of the coiled-coil region, but it lacks the serine-rich and acidic domains and the nuclear-localization sequence. The mechanism of dominant-negative regulation is unclear, but as the coiled-coil region in the mutant may enable PML to homodimerize¹⁵, the mutant protein may dimerize with wild-type PML and thus interfere with its function.

PML is expressed in multiple isoforms¹⁶. To determine whether the PML-F12 isoform is expressed in normal mouse cells, we

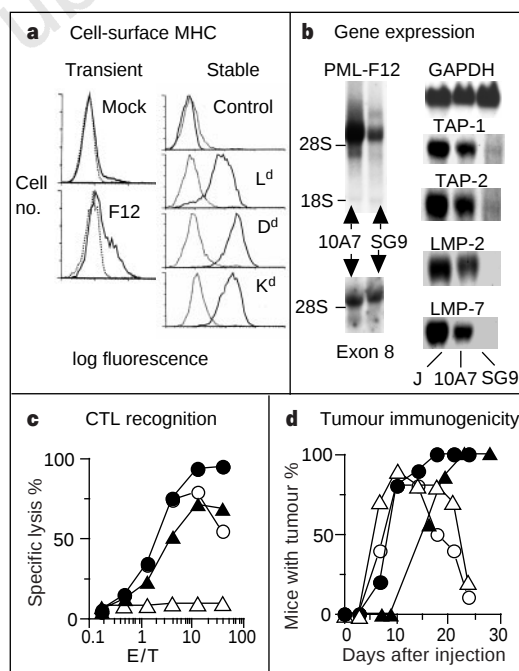


Figure 2 Identification by expression cloning of a cDNA that compensates for the antigen-presentation defects of the ReB7 cell line. **a**, Restoration of cell-surface H-2L^d expression by a single cDNA clone. Left, F12 induces MHC class I-L^d in transient spheroplast fusion assays. Top, cell-surface L^d in ReB7 fused with spheroplasts from bacteria transfected with pCDM8 vector alone; bottom, L^d expressed on ReB7 fused with F12 spheroplasts. Dotted lines represent binding of second-step reagent only; solid lines depict binding of anti-L^d monoclonal antibody HB27. Right, comparison of cell-surface MHC class I on ReB7 transfected with either vector EBO-Sfi alone (SG9, dotted lines), or F12 insert (10A7, solid lines). **b**, Expression of PML (left), or TAP-1, TAP-2, LMP-2, and LMP-7 (right) in stable clones transfected with either PML-F12 (10A7) or EBO-Sfi vector alone (SG9) or in untransfected J558-B7 (J). PML expression as measured by full-length PML-F12 probe (top), or PML exon 8 (bottom). **c**, PML-F12 cDNA restores CTL recognition of ReB7 tumour: cytotoxicity of SG9 (triangles) or 10A7 (circles) in the presence (black symbols) or absence (white symbols) of exogenous P1A peptide. **d**, Tumorigenicity of parental J558-B7 (white circles), ReB7 (black circles), 10A7 (white triangles), and SG9 (black triangles). Syngeneic BALB/c mice were injected with 5 $\times$ 10⁶ tumour cells; the incidence of tumours was determined by physical examination.

isolated RNA from normal spleen and thymus and analysed it for the isoforms of PML by RT-PCR by using a forward primer on exon 4 and a reverse primer on exon 9. Two additional PML cDNA isoforms, one containing all exons and the other lacking exon 8, were cloned from J558-B7 and used as controls. RT-PCR amplified three major PML bands from normal spleen and thymus (Fig. 4a): the top band corresponds to PML containing all exons; the second to PML lacking either exon 5 (E5⁻) or exon 8 (E8⁻) (indistinguishable because of their similar sizes of 509 or 521 base pairs); and the third to the PML-F12 isoform that lacks both exons 5 and 8. We cloned the third band and sequenced three individual clones, which confirmed that all clones were identical to PML-F12, indicating that PML-F12 is a major isoform in normal cells. To test whether PML controls cell-surface MHC class I expression, we transfected

NIH3T3 cells with vector, PML-F12, or PML-E5⁺E8⁻. NIH3T3 cells derived from Stat-1-deficient mice expressed no detectable cell-surface MHC class I. Transfection of PML-F12 induced the expression of significant amounts of cell-surface MHC class I, whereas transfection with a PML-E5⁺E8⁻ clone did not upregulate cell-surface MHC (Fig. 4b). We conclude that not all PML isoforms can control the expression of cell-surface MHC. PML-F12 also increased cell-surface MHC in wild-type 3T3 cells (data not shown), so PML-F12 must induce MHC class I antigen presentation in normal cells.

As PML has features of a transcription factor, we tested whether PML-F12 could be a transactivator for antigen-processing genes. We have recently cloned a 4-kilobase (kb) genomic DNA fragment upstream of the TAP-2 coding region and found that a 1.7-kb, 5'

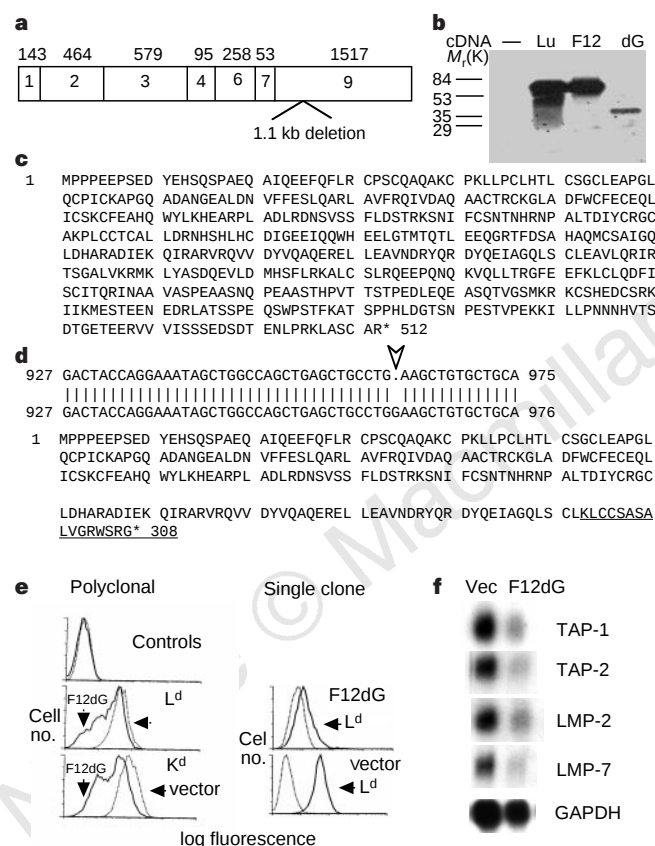


Figure 3 Identification of F12 cDNA as a normal isoform of the PML gene and of a dominant-negative mutation of PML in ReB7. **a**, PML exons in F12 cDNA (sizes in base pairs of each exon are shown at the top). **b**, *In vitro* translation using no cDNA (-), or control luciferase cDNA (Lu), PML-F12 (F12), or PML-F12dG (dG). **c**, Predicted amino-acid sequence of F12 cDNA. Asterisk indicates stop. **d**, Frameshift mutation of the PML gene in the ReB7 cell line. Top, partial sequence of mutated (mutation present in 2/6 of the clones sequenced; arrowhead), and unmutated (4/6 clones) PML gene in the same subline of ReB7. Five clones from J558-B7 did not contain this mutation. Bottom, predicted amino-acid sequence of the protein encoded by the mutated PML gene. **e**, PML mutant encodes a dominant-negative regulator for expression of cell-surface MHC class I. Left, J558-B7 cell line was transfected with either EBO-Sfi vector alone (dotted lines), or the mutant PML gene (F12dG, solid lines). After selection with hygromycin, drug-resistant cells were stained with second-step reagent alone (top) or monoclonal antibodies specific for L^d (HB27, middle) or K^d (HB159, bottom). Right, representative clone from J558-B7 transfected with either vector alone (top) or PML-F12dG (bottom). **f**, Downregulation of TAP-1, TAP-2, LMP-2 and LMP-7 by PML-F12dG.

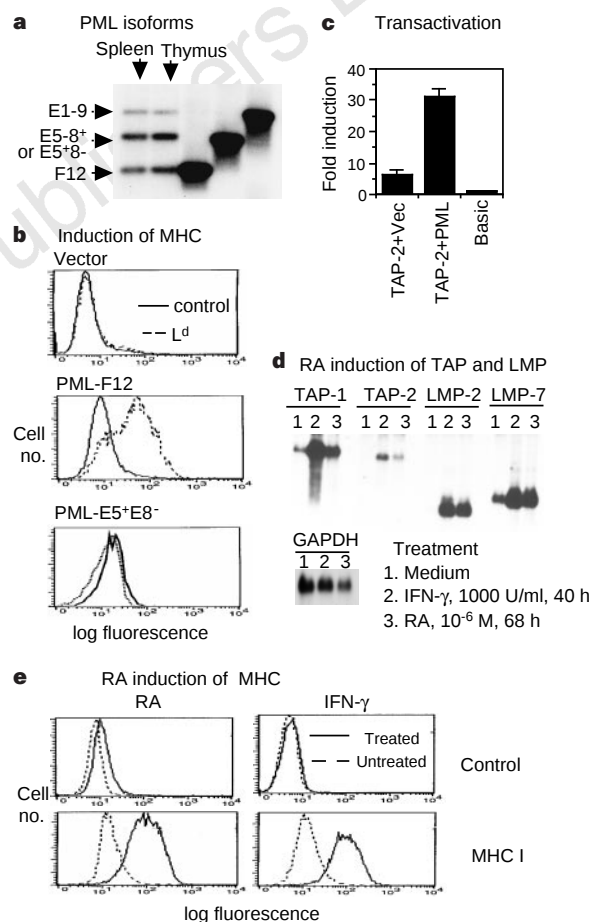


Figure 4 Expression and function of PML-F12 isoform in normal murine cells. **a**, Expression of PML isoforms, as determined by RT-PCR using a forward primer on exon 4 and a reverse primer on exon 9. As controls, we used PML-F12 (F12, lane 3), PML lacking exon 8 (E5⁻E8⁻, lane 4), or PML consisting of all exons (E1-9, lane 5) as templates. Note that normal spleen and thymus have 3 or 4 major forms of PML. E1-9 (top band, 660 bp), E5⁻E8⁺ or E5⁺E8⁻ (middle band, size 509 or 521 bp), and F12 (bottom band, 370 bp). **b**, PML-F12, but not PML-E5⁺E8⁻, upregulates cell-surface MHC in murine embryonic fibroblast cells. 3T3 cells prepared from Stat-1(-/-) C57BL/6j mice were transfected with either vector alone (top), or PML-F12 cDNA (middle) or PML-E5⁺E8⁻ cDNA (bottom). Results are representative of one of 6-15 clones from each group. **c**, Transactivation of TAP-2 promoter by PML-F12. NIH3T3 cells were transfected with either pGL2 (basic), pGL2-1.7 kb TAP-2 promoter+EBP-Sfi-vector (TAP-2+Vec), or pGL2-1.7 kb TAP-2 promoter+EBP-PML-F12 (TAP-2+PML). **d**, **e**, Effects of all-trans retinoic acid (RA) and interferon (IFN- γ) on the expression of antigen-processing genes (**d**) and cell-surface MHC class I in NB4 cell lines (**e**) using a FITC-labelled, anti-HLA-A, B, C monoclonal antibody from Pharmingen.

flanking fragment contains a promoter for TAP-2 (Y.G. *et al.*, manuscript in preparation). We therefore co-transfected PML-F12 with the 1.7-kb TAP-2 promoter into NIH3T3 cells and found that PML-F12 significantly increased expression of the TAP-2 promoter (Fig. 4c).

The PML gene was first identified as a fusion partner of the retinoic acid receptor- α (RAR- α) gene that is translocated in acute promyelocytic leukaemia (APL) cells and acts as a dominant oncogene^{17–20}. As shown in Fig. 4d, e, the human APL cell line NB4 has barely any detectable TAP-1, TAP-2, LMP-2, LMP-7 or cell-surface MHC class I. More important, treatment of NB4 with all-*trans* retinoic acid, which causes degradation of the PML–RAR α fusion protein²¹ and differentiation of NB4 cells, triggers the expression of antigen-presentation genes and cell-surface MHC. These results indicate that malfunction of PML can affect recognition by T cells of acute promyelocytic leukaemia and may allow tumours to escape from the host immune system.

Taken together, our results show that overexpression of PML compensates for antigen-presentation defects in untransformed cells as well as in recurrent tumours carrying a dominant-negative mutation in exon 3 of the PML gene. As PML is a primary target of interferon²² and acts downstream of Stat-1, it may be involved in the interferon-mediated induction of TAP-1, TAP-2, LMP-2, LMP-7 and MHC class I heavy chains^{1–5,23}. The relation between PML and the IRF-1, which also acts downstream of Stat-1 and is involved in the control of antigen presentation¹, is still unclear. In tissues with low levels of MHC class I, such as the brain, sperm and placenta, the amount of PML is also low²⁴. Moreover, a significant proportion of tumour tissues, particularly those that are more malignant, show reduced expression of cell-surface MHC class I owing to multiple antigen-presentation defects^{8–10}. Downregulation of PML occurs in invasive and metastatic lesions of a large collection of tumours^{24,25}. Finally, adenoviral proteins, which downregulate the transcription of genes devoted to MHC class I antigen presentation^{6,7}, are associated with a PML domain and affect the nuclear distribution of PML²⁶, possibly disrupting its function. The general significance of PML malfunction in tumour evasion of host immunity remains to be determined. □

Methods

Cell lines. The tumour cell line J558 was supplied by ATCC. B7-1-transfected J558 cells (J558-B7) have been described¹¹. ReB7, a spontaneous tumour revertant, was isolated from a syngeneic mouse that had been injected with a J558-B7 tumour three months before. The mouse rejected the tumour within three weeks, but it recurred two months after the injection. Single-cell suspensions were prepared from each tumour. Most tumour cells recovered from the immune-competent mice expressed a low level of MHC class I; a few class I MHC^{high} cells were removed by two rounds of FACS sorting. Murine 3T3 cells deficient for Stat-1(–/–) were derived from Stat-1(–/–) mice²⁷. Human promyelocytic NB4 cells²⁸ were provided by M. Lanotte.

cDNA complementation and expression cloning. A cDNA library prepared from RAW8.1 (ref. 29) was introduced into ReB7 by spheroplast fusion³⁰. Viable cells were isolated at 16 h after fusion and incubated with anti-H-2L^d mAb HB27. After washing off unbound antibody, L^d-expressing cells were enriched by panning onto a plate precoated with Fab fragment of rabbit anti-mouse IgG (Accurate Chemical). After 2–3 h incubation, unbound cells were removed by four gentle washes. Episomal DNA was isolated from adherent cells and the plasmid was amplified in *E. coli* MC1061/p3. Pools of ~40 clones each were used to transfect ReB7 by spheroplast fusion. Expression of H-2L^d from each pool was determined by flow cytometry. Positive pools were subdivided and spheroplast fusion was repeated. Individual clones from a positive pool were tested again and one clone (F12) was picked because it induced significant cell-surface H-2L^d expression.

Northern-blot analysis. Poly(A)⁺ RNA, prepared from total RNA using a Micro Poly(A) Pure kit (Ambion), was used for northern blotting. DNA templates for labelling were obtained by cloning murine and human TAP-1, TAP-2, LMP-2 and LMP-7 from RT-PCR products and their identities

confirmed by DNA sequencing. Full-length K^d cDNA was provided by S. Vukamnovic.

Transactivation assay of TAP-2 promoter. A 4-kb genomic DNA was amplified by PCR using primers corresponding to the 3' end of LMP-7 and the 5' end of the TAP-2 coding region; its identity was confirmed by sequencing. A 1.7-kb fragment upstream of the TAP-2 coding region that had maximal promoter activity was used. Transactivator activity of the PML gene was measured using the dual-luciferase reporter assay system (Promega) according to the manufacturer's instructions. All plasmids were used at 1 μ g per well in a 24-well plate. pRL-SV40 control DNA was used to normalize transfection efficiency. Data are presented as fold induction relative to cells transfected with vector only, after normalizing the transfection efficiency. Results are presented as means and standard errors of quadruplicate experiments.

Received 14 September; accepted 13 October 1998.

- White, L. C. *et al.* Regulation of LMP2 and TAP1 genes by IRF-1 explains the paucity of CD8 T cells in IRF-1(–/–) mice. *Immunity* **5**, 365–376 (1996).
- Matinez, C. K. & Monaco, J. J. Homology of proteasome subunits to a major histocompatibility complex-linked LMP gene. *Nature* **353**, 554–567 (1991).
- Trowsdale, J. *et al.* Sequences encoded in the class II region of the MHC related to the 'ABC' superfamily of transporters. *Nature* **348**, 741–744 (1990).
- Townsend, A. R. M. *et al.* The epitope of influenza nucleoprotein recognized by cytotoxic lymphocytes can be defined with short synthetic peptides. *Cell* **44**, 959–968 (1986).
- Spies, T. *et al.* A gene in the human major histocompatibility complex class II region controlling the class I antigen-presentation pathway. *Nature* **348**, 744–747 (1990).
- Rotem-Yehudar, R., Winograd, S., Sela, S., Coligan, J. E. & Ehrlich, R. Downregulation of peptide transporter genes in cell lines transformed with highly oncogenic adenovirus 12. *J. Exp. Med.* **180**, 477–488 (1994).
- Rotem-Yehudar, R., Groettrup, M., Soza, A., Klotzel, P. M. & Ehrlich, R. LMP-associated proteolytic activities and TAP-dependent peptide transport for class I MHC molecules are suppressed in cell lines transformed by the highly oncogenic adenovirus 12. *J. Exp. Med.* **183**, 499–514 (1996).
- Restifo, N. P. *et al.* Identification of human cancers deficient in antigen-processing. *J. Exp. Med.* **177**, 265–272 (1993).
- Korkolopoulou, P., Kaklamanis, L., Pezzella, F., Harris, A. L. & Gatter, K. C. Loss of antigen-presenting molecule (MHC class I and TAP-1) in lung cancer. *Br. J. Cancer* **73**, 148–153 (1996).
- Vegh, Z., Wang, P., Yank, F. & Klein, E. Selective down-regulated expression of major histocompatibility complex class I alleles in human solid tumors. *Can. Res.* **53**, 2416–2420 (1993).
- Ramarathinam, L., Castle, M., Wu, Y. & Liu, Y. T. Cell costimulation by B7/BB1 induces CD8 T cell-dependent tumor rejection: an important role of B7/BB1 in the induction, recruitment, and effector function of antitumor T cells. *J. Exp. Med.* **179**, 1205–1214 (1994).
- Ramarathinam, L. *et al.* Multiple lineages of tumors express a common tumor antigen, P1A, but they are not cross-protected. *J. Immunol.* **155**, 5323–5329 (1995).
- Steimle, V., Otten, L. A., Zufferey, M. & Mach, B. Complementation cloning of an MHC class II transactivator mutated in hereditary MHC class II deficiency (or bar lymphocyte syndrome). *Cell* **75**, 135–146 (1993).
- Goddard, A. D. *et al.* Cloning of the murine homolog of the leukemia-associated PML gene. *Mamm. Gen.* **7**, 732–737 (1995).
- Perez, A. *et al.* PML/RAR homodimers: distinct DNA binding properties and heterodimeric interaction with RXR. *EMBO J.* **12**, 3171–3182 (1993).
- Fagioli, M. *et al.* Alternative splicing of PML transcripts predicts coexpression of several carboxy-terminally different protein isoforms. *Oncogene* **7**, 1083–1091 (1992).
- de Thé, H. C. *et al.* The PML–RAR α fusion mRNA generated by the t(15;17) translocation in acute promyelocytic leukemia encodes a functionally altered RAR. *Cell* **66**, 675–684 (1991).
- Goddard, A. D., Borrow, J., Freemont, P. S. & Solomon, E. Characterization of a zinc finger gene disrupted by the t(15;17) in acute promyelocytic leukemia. *Science* **254**, 1371–1374 (1991).
- Grignani, F. *et al.* The acute promyelocytic leukemia-specific PML–RAR α fusion protein inhibits differentiation and promotes survival of myeloid precursor cells. *Cell* **74**, 423–431 (1993).
- Kakizuka, A. *et al.* Chromosomal translocation t(15;17) in human acute promyelocytic leukemia fuses RAR α with a novel putative transcriptional factor, PML. *Cell* **66**, 663–674 (1991).
- Yoshida, H. *et al.* Accelerated degradation of PML-retinoic acid receptor alpha (PML–RAR α) oncoprotein by all-*trans*-retinoic acid in acute promyelocytic leukemia: possible role of the proteasome pathway. *Cancer Res.* **56**, 2945–2948 (1996).
- Lavau, C. *et al.* The acute promyelocytic leukemia-associated PML gene is induced by interferon. *Oncogene* **11**, 871–876 (1995).
- Blonar, M. A., Baldwin, A. S. Jr, Flavell, R. A. & Sharp, P. A. A γ -interferon-induced factor that binds the interferon response sequence of MHC class I gene, H-2K^b. *EMBO J.* **8**, 1139–1144 (1989).
- Gambacorta, M. *et al.* Heterogeneous nuclear expression of the promyelocytic leukemia (PML) protein in normal and neoplastic human tissue. *Am. J. Pathol.* **149**, 2023–2035 (1996).
- Koken, M. H. M. *et al.* The PML growth-suppressor has an altered expression in human oncogenesis. *Oncogene* **10**, 1315–1324 (1995).
- Carvalho, T. *et al.* Targeting of adenovirus E1A and E4-ORF3 proteins to nuclear matrix-associated PML bodies. *J. Cell. Biol.* **131**, 45–56 (1995).
- Durbin, J. E., Hackenmiller, R., Simon, M. C. & Levy, D. E. Targeted disruption of the mouse stat1 gene results in compromised innate immunity to viral disease. *Cell* **84**, 443–450 (1996).
- Lanotte, M. *et al.* NB4, a maturation-inducible cell line with t(15;17) marker isolated from a human promyelocytic leukemia (M3). *Blood* **77**, 1080–1086 (1991).
- Guo, Y., Wu, Y., Shinde, S. & Liu, Y. Identification of CD44H as a costimulatory molecule rapidly induced by CD40 ligand. *J. Exp. Med.* **184**, 955–961 (1996).
- Oi, V. T., Morrison, S. L., Herzenberg, L. A. & Berg, P. Immunoglobulin gene expression in lymphoid cells. *Proc. Natl Acad. Sci. USA* **80**, 825–829 (1983).

Acknowledgements. We thank V. Nussenzweig, X.-H. Sun and X. D. Fu for discussions and for critically reading the manuscript; J. Hirst for flow cytometry; B. Mach for the EBO-Sf vector; S. Vukamnovic for the K^d cDNA clone; M. Lanotte for NB4 cells; and F. Hitchcock for help with preparing the manuscript. This study was supported by grants from the NIH and by the Kaplan Comprehensive Center of NYU Medical Center and by Ohio State University Comprehensive Cancer Center.

Correspondence and requests for materials should be addressed to Y.L. (e-mail: liu-3@medctr.osu.edu).

Interdomain communication regulating ligand binding by PPAR- γ

Dalei Shao*, Shamina M. Rangwala*, Shannon T. Bailey*, Samuel L. Krakow*, Mauricio J. Reginato† & Mitchell A. Lazar*‡

* Division of Endocrinology, Diabetes, and Metabolism, Department of Medicine, and Departments of † Pharmacology and ‡ Genetics, University of Pennsylvania School of Medicine, Philadelphia, Pennsylvania 19104, USA

Binding to receptors in the cell nucleus is crucial for the action of lipophilic hormones and ligands. PPAR- γ (for peroxisome proliferator-activated receptor) is a nuclear hormone receptor that mediates adipocyte differentiation^{1,2} and modulates insulin sensitivity³, cell proliferation⁴ and inflammatory processes^{5,6}. PPAR- γ ligands have been implicated in the development of atherogenic foam cells⁷ and as potential cancer treatments⁸. Transcriptional activity of PPAR- γ is induced by binding diverse ligands, including natural fatty acid derivatives^{9–11}, antidiabetic thiazolidinediones¹², and non-steroidal anti-inflammatory drugs¹³. Ligand binding by PPAR- γ , as well as by the entire nuclear-receptor superfamily, is an independent property of the carboxy-terminal ligand-binding domain (LBD) of the receptor^{14,15}. Here we show that ligand binding by PPAR- γ is regulated by

intramolecular communication between its amino-terminal A/B domain and its carboxy-terminal LBD. Modification of the A/B domain, for example by physiological phosphorylation by MAP kinase, reduces ligand-binding affinity, thus negatively regulating the transcriptional and biological functions of PPAR- γ . The ability of the A/B domain to regulate ligand binding has important implications for the evaluation and mechanism of action of potentially therapeutic ligands that bind PPAR- γ and that are likely to extend to other members of the nuclear-receptor superfamily.

Extracellular signals leading to MAP kinase activation cause serine 112 in the A/B domain to be phosphorylated, thereby inhibiting PPAR- γ ^{16–18}, but the molecular mechanism of this process is unknown. Mutation of serine 112 to aspartate (S112D) reduces the electrophoretic mobility of PPAR- γ 2 by much the same amount as does phorbol ester-induced phosphorylation of wild-type PPAR- γ at this site (Fig. 1a). The high-affinity thiazolidinedione (TZD) ligand, BRL49653, was nearly 10-fold more effective, as measured by the half-maximal effective concentration (EC₅₀), in transcriptional activation of PPAR γ 2-S112A than it was for the S112D mutant (Fig. 1b). That this was an active effect of the A/B domain was shown by the potency of BRL49653 with PPAR γ 2 Δ N (lacking amino acids 1–118 of the A/B domain) being equal to that with PPAR γ 2-S112A (data not shown). Co-expression of the isolated A/B domain of S112D (in *trans*) had no effect on the EC₅₀ of BRL49653 with the Δ N, S112A or S112D mutants (data not shown). We did not detect interaction between any of the A/B domains and the LBD of PPAR- γ 2 in yeast and mammalian two-hybrid assays, or in *in vitro* glutathione S-transferase (GST) pulldown assays (data not shown).

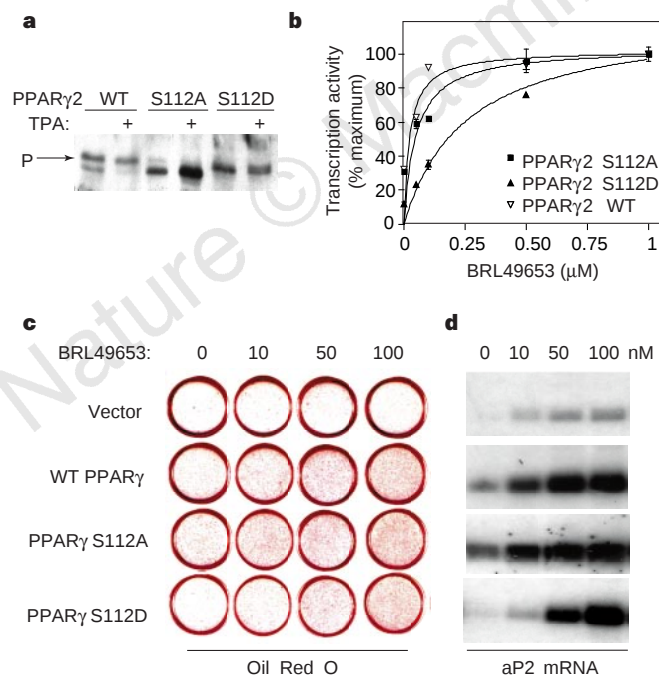


Figure 1 Potency of BRL49653 depends upon the PPAR- γ A/B domain. **a**, Alterations in PPAR- γ A/B domain affect its gel mobility. Wild-type and mutant PPAR- γ were stably expressed in 3T3-L1 preadipocytes using retroviral vectors. TPA, 12-O-tetradecanoylphorbol-13-acetate; P, phosphorylated band; WT, wild type. **b**, Transcriptional potency of BRL49653 depends upon the PPAR- γ A/B domain. (Acyl-CoA oxidase PP2E $\times$ 3)-TK-luciferase reporter gene was transfected into 293T cells with the indicated PPAR- γ mutants. Equal expression of the PPAR- γ constructs was confirmed by western blot (results not shown). **c**, **d**, Adipogenic potency of BRL49653 depends upon the PPAR- γ A/B domain. Wild-type and mutant PPAR- γ were stably expressed in 3T3-L1 cells using retroviral vectors. Oil-red O staining (**c**) and northern blot analysis for aP2 expression (**d**) are shown. Equal loading of RNA was confirmed by staining for ribosomal RNA (results not shown).

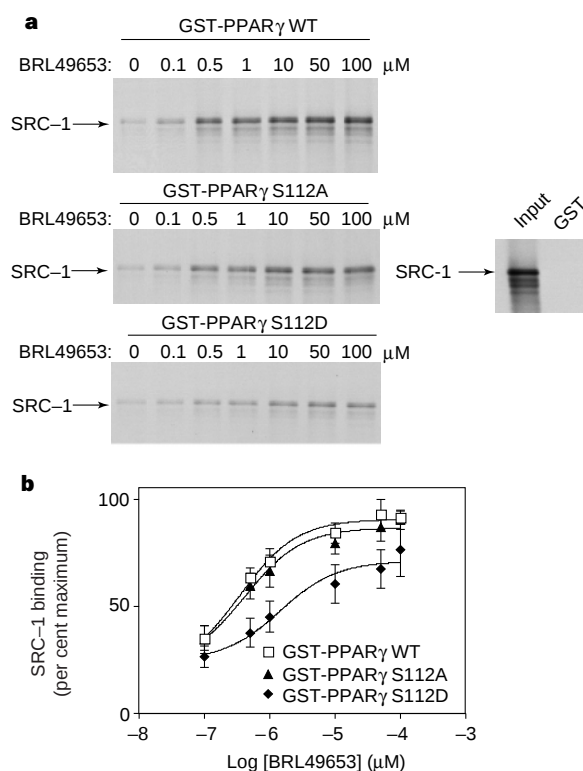


Figure 2 Potency of BRL49653 in co-activator recruitment to PPAR- γ LBD depends upon the A/B domain. **a**, Ligand-dependent interaction of SRC-1 with GST-PPAR γ 2 or GST. **b**, Phosphorimager quantification of the three experiments shown in **a**.

The effects of the A/B domain on adipogenicity of PPAR- γ 2 were similar to the transcriptional results. The effect of BRL49653 was five to ten times more potent on differentiating 3T3-L1 cells which ectopically express PPAR γ 2-S112A rather than PPAR γ 2-S112D, as assessed by Oil-red O staining of accumulated lipid (Fig. 1c), as well as by expression of the PPAR-responsive adipocyte-specific gene *aP2* (Fig. 1d). These results are in agreement with the potency of the wild-type and S112A forms of PPAR- γ 2 in NIH 3T3 fibroblasts¹⁸.

We next examined whether A/B domain-dependent differences in potency of BRL49653 are reflected in ligand-dependent recruitment of the nuclear hormone receptor (NHR) co-activator protein SRC-1. Figure 2a shows that in the absence of ligand, wild-type PPAR- γ 2 fused to GST interacts weakly with SRC-1. Addition of BRL49653 led to a dose-dependent increase in SRC-1 binding. The EC₅₀ of BRL49653 for increased SRC-1 interaction was similar for the wild type and the S112A mutant, between 100 and 500 nM, which is consistent with the reported *K_d* for BRL49653. By contrast, although the S112D mutant also interacts with SRC-1, the EC₅₀ of BRL49653 was nearly 10-fold greater than for the other forms of PPAR- γ 2. The results of several experiments are plotted in Fig. 2b. In addition to the reduced potency of BRL49653 with respect to the S112D mutant, there was a modest reduction in maximal SRC-1 binding. It is possible that the wild-type and S112A A/B domains contribute to the interaction of PPAR- γ with SRC-1.

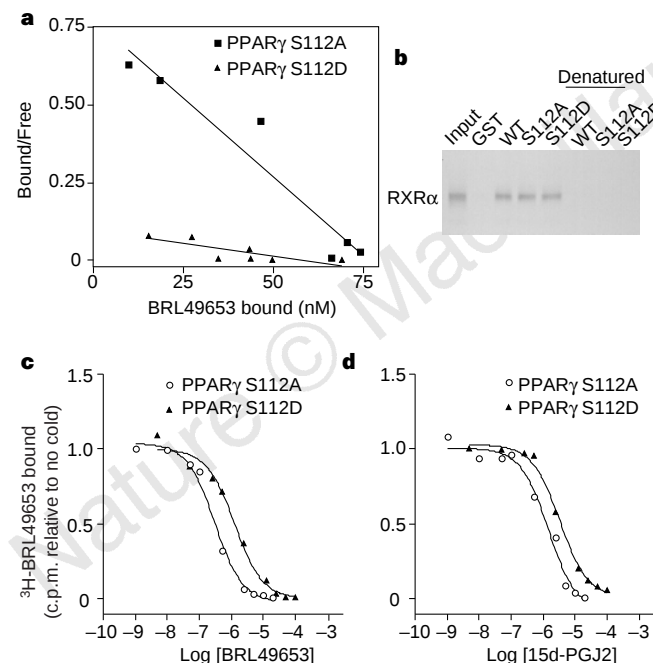


Figure 3 Affinity of BRL49653 for PPAR- γ depends upon the A/B domain. **a**, Direct measurement of affinity by Scatchard analysis. **b**, RXR interaction with PPAR- γ is independent of the A/B domain. Pull-down of ³⁵S-RXR with the GST-PPAR γ proteins. Purity was checked and the amounts of GST proteins were confirmed as being equal by Coomassie blue staining (results not shown). **c**, Displacement of [³H]BRL49653 from GST-PPAR γ 2 proteins by cold BRL49653. **d**, Displacement of [³H]BRL49653 from GST-PPAR γ 2 proteins by cold 15d-PGJ2.

However, we found no interaction between the isolated A/B domains and SRC-1 in GST pulldown assays (data not shown).

The reduced potency of BRL49653 with respect to the S112D mutant in transactivation, adipogenesis and co-activator recruitment suggests that ligand binding of PPAR- γ is impaired in the S112D mutant, which we confirmed by a Scatchard plot of BRL49653 binding (Fig. 3a). PPAR γ 2-S112A had a binding constant for BRL49653 of 99 nM, consistent with the EC₅₀ for SRC-1 recruitment and also with what we (data not shown) and others^{10,11} have found for the LBD in the absence of the A/B and DNA-binding domains. In contrast, the affinity of PPAR γ 2-S112D for BRL49653 was approximately 6-fold lower, at 622 nM. The S112A and S112D proteins were comparable in function, including their interaction with the retinoid-X receptor (RXR), which was dependent on correct folding of the proteins (Fig. 3b). Figure 3c shows a ligand-displacement assay which confirms the affinity difference between PPAR γ 2-S112A and S112D for BRL49653; the mean affinity difference in two experiments was 5.1-fold. The affinity of a structurally different prostaglandin ligand, 15 deoxy- Δ 12,14-PGJ2 (15d-PGJ2), was also reduced for S112D relative to S112A (Fig. 3d). This reduction was less marked (mean difference in two experiments was 3.1-fold), which might contribute to cell-specific differences in the potencies of TZDs and 15d-PGJ2 (refs 5, 6).

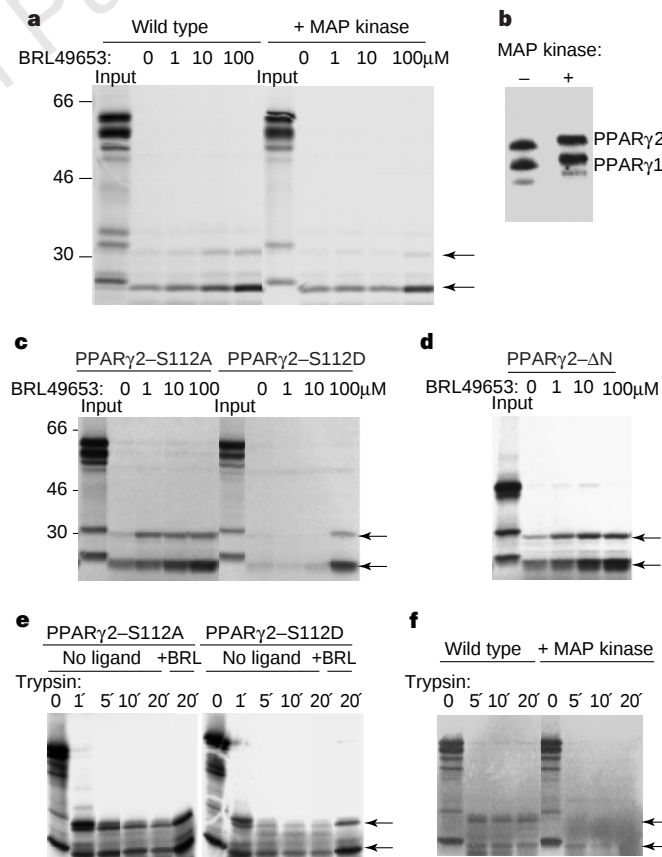


Figure 4 The A/B domain regulates potency of BRL49653 in protease protection and the conformation of unliganded PPAR- γ . **a**, Dose response for BRL49653 in a protease protection assay of non-phosphorylated versus MAP kinase-phosphorylated PPAR- γ 2. **b**, Quantitative phosphorylation of PPAR- γ 2 treated with MAP kinase *in vitro*. **c**, Protease protection of PPAR γ 2-S112A versus S112D by BRL49653. The input doublet is due to internal initiation of translation producing PPAR- γ 1. **d**, Protease protection of PPAR γ 2- Δ N by BRL49653. Arrows point to protected bands, which were not recognized by N-terminal antibody (results not shown). **e**, Limited tryptic proteolysis of PPAR γ 2-S112A and S112D in the absence of ligand. **f**, Limited tryptic proteolysis of non-phosphorylated versus MAP kinase-phosphorylated PPAR- γ 2 in the absence of ligand.

Ligand binding by NHRs induces a conformational change which results in protection of the C terminus from proteolysis¹⁹. Figure 4a shows that a C-terminal fragment of wild-type PPAR- γ 2 is protected from tryptic proteolysis by BRL49653, consistent with previous reports²⁰. The protection was dose-dependent, although the EC₅₀ for BRL49653 in this assay was higher than the K_d, presumably because proteolysis is irreversible, which prevents true equilibrium measurements, especially given that the half-time for dissociation of BRL49653 from PPAR- γ is less than 20 min (data not shown). When PPAR- γ 2 was phosphorylated *in vitro* with recombinant activated MAP kinase (as confirmed by the SDS-PAGE migration shift shown in Fig. 4b), the EC₅₀ for BRL49653 in this assay shifted to the right, consistent with an almost log-order reduction in affinity. In support of the functional similarity between phosphorylated PPAR- γ and its S112D mutant, results were comparable when the S112A and S112D mutants were compared (Fig. 4c), showing the EC₅₀ for the increase in the ligand-protected band by BRL49653 to be ~10-fold higher for S112D. Figure 4d shows that the inhibitory effect of the negatively charged A/B domain was dominant, because the effects of BRL49653 on PPAR γ 2 Δ N were similar to those on the wild type or on the S112A mutant.

In all cases, the pattern of protected bands was the same at saturating concentrations of ligand, indicating that the conformations of the liganded receptors could be similar. This suggests that the mechanism by which the N terminus modulates ligand binding is by altering the conformation of the unliganded receptor. No three-dimensional structure of full-length PPAR- γ or of any NHR is available. We therefore investigated the mechanism of the N-terminal effects on ligand binding. Circular dichroism of the unliganded S112A and S112D proteins (performed in collaboration with R. Marmorstein) revealed no differences in secondary structure between the two (data not shown). By contrast, their tertiary structures, as probed by limited protease digestion of the unliganded receptors, were under the influence of the N terminus (Fig. 4e). A proteolytic fragment that was more resistant to tryptic degradation in the S112A than in the S112D form co-migrated with the band protected by ligand from harsher proteolytic conditions. More dramatic results were obtained by comparing the native and MAP kinase-phosphorylated forms of PPAR- γ 2 (Fig. 4f), suggesting that the conformation of unliganded unphosphorylated PPAR- γ , or of the S112A mutant, is closer to that of the liganded form, explaining the increased affinity for ligand and of phosphorylated PPAR- γ for the co-repressor SMRT²¹, which recognizes unliganded NHR. We believe that the effect of the A/B domain is intramolecular as PPAR- γ is primarily a monomer under these solution conditions (data not shown), but intermolecular interactions may play a part. Thus, the N terminus regulates PPAR- γ ligand binding by altering the conformation of the unliganded receptor, and hence the free energy of the transition to the liganded conformation.

The ability of the A/B domain of PPAR- γ to alter ligand binding affinity by nearly a log order has important implications for diabetes. LBD mutations in thyroid receptors that reduce affinity by this amount result in clinical syndromes of hormone resistance²². Our results indicate that treatments causing reduced phosphorylation of PPAR- γ *in vivo* could improve the potency of both TZDs and endogenous ligands, perhaps related to 15d-PGJ₂. Growth factors, tumour promoters and prostaglandins all regulate PPAR- γ phosphorylation^{16,18,23}; their effects may vary with cell type and during the cell cycle, thereby affecting the affinity of PPAR- γ ligands that have been considered as therapeutic for cancer, vascular and inflammatory diseases.

The A/B domain of PPAR- γ , including S112, is highly conserved, as in other NHR subtypes, although the function of the terminal A/B domain of NHRs is unclear. The A/B domains of retinoic acid receptor related orphan receptor- α (ROR- α) and thyroid hormone receptor- α regulate DNA-binding affinity²⁴, in the latter case in a phosphorylation-dependent way²⁵. The A/B domain of some NHRs

contains a ligand-independent transactivation domain (AF-1) which, in the case of the oestrogen receptor, can be regulated by MAP kinase phosphorylation^{26,27}, and which seems to function by binding to the LBD through SRC-1 (ref. 28). We have now demonstrated how the A/B domain of NHRs can function as a target for extracellular signals to regulate ligand-binding affinity, providing a new mechanism for crosstalk through regulation of ligand activation of this important class of signalling molecules. □

Methods

3T3-L1 cell culture and differentiation. 3T3 L1 cells were cultured in growth medium containing 10% fetal calf serum in high-glucose DMEM. BRL49653 was added to post-confluent cells for 10 days. PPAR- γ 2 wild-type and S112A mutants were constructed and used as described², and the S112D mutant was created by using PCR and verified by sequencing. Oil-red O staining and northern blot analysis were done as described².

Transient transfections. 293T cells were cultured in 10% fetal bovine serum. Before transfection, cells were switched to medium containing charcoal-stripped serum. Ligands were added 16 h after transfection, and cells were collected 24 h later. Results were normalized to expression of a co-transfected β -galactosidase-expressing plasmid, as described²⁹. Western blot analysis was done after electrophoresis of 50 μ g 3T3-L1 protein extract in 10% polyacrylamide gels containing 4 M urea²³.

Co-activator recruitment assay. This was done as previously described³⁰, with minor modifications. GST-fusion proteins bound to glutathione-Sepharose beads (100 μ l of a 50% slurry) in binding buffer (50 mM KCl, 20 mM HEPES, pH 7.9, 2 mM EDTA, 0.1% Nonidet P-40, 10% glycerol, 0.5% non-fat dry milk, 5 mM dithiothreitol (DTT)) were mixed with 5 μ l radiolabelled *in vitro* translated SRC-1 and ligand or vehicle, incubated at 4 °C for 1 h, and washed with binding buffer. Bound SRC-1 was analysed by SDS-PAGE.

Ligand-binding assays. Bacterially expressed GST-PPAR γ 2 fusion proteins purified on glutathione-Sepharose beads (Pharmacia) were incubated with [³H]BRL49653 (18 Ci mmol⁻¹) in binding buffer (10 mM Tris, pH 8.0, 50 mM KCl, 10 mM DTT). Following incubation for 2 h at 4 °C, beads were washed 5 times with binding buffer, and bound ³H was determined by liquid scintillation counting. Each experiment was repeated at least three times with similar results. For Scatchard analysis, 25 nM [³H]BRL49653 was used, either alone or with increasing concentrations of cold ligand. For displacement assays, 100 nM [³H]BRL49653 was used. Competitor ligand or solvent carrier ethanol was added as indicated.

Protease-sensitivity assay. This was performed as previously described²⁰, with minor modifications. Proteins were translated *in vitro* in the presence of ³⁵S-radiolabelled methionine using the TNT kit (Promega). When applicable, purified MAP kinase (New England BioLabs) was used as directed by the manufacturer. The protein was incubated with 1 μ l BSA (1 mg ml⁻¹) and 1 μ l vehicle or ligand in a final volume of 10 μ l at 25 °C for 20 min, after which 0.5 μ l PBS or trypsin (1 mg ml⁻¹) was added to a 4.5- μ l aliquot. After 20 min (or varying times as indicated) at 25 °C, 1 μ l soybean trypsin inhibitor (5 mg ml⁻¹) was added, and proteins were analysed by SDS-PAGE (12% polyacrylamide gel).

Received 13 August; accepted 28 September 1998.

1. Tontonoz, P., Hu, E. & Spiegelman, B. M. Stimulation of adipogenesis in fibroblasts by PPAR γ 2, a lipid-activated transcription factor. *Cell* **79**, 1147–1159 (1994).
2. Shao, D. & Lazar, M. A. PPAR γ , C/EBP α , cell cycle status and the commitment to adipocyte differentiation. *J. Biol. Chem.* **272**, 21473–21478 (1997).
3. Willson, T. M. *et al.* The structure–activity relationship between peroxisome proliferator-activated receptor γ and the antihyperglycemic activity of thiazolidinediones. *J. Med. Chem.* **39**, 665–668 (1996).
4. Altio, S., Xu, M. & Spiegelman, B. M. PPAR γ induces cell cycle withdrawal: inhibition of E2F/DP DNA-binding activity via down-regulation of PP2A. *Genes Dev.* **11**, 1987–1998 (1997).
5. Jiang, C., Tring, A. T. & Seed, B. PPAR- γ agonists inhibit productions of monocyte inflammatory cytokines. *Nature* **391**, 82–86 (1998).
6. Ricote, M., Li, A. C., Willson, T. M., Kelly, C. J. & Glass, C. K. The peroxisome proliferator-activated receptor- γ is a negative regulator of macrophage activation. *Nature* **391**, 79–82 (1998).
7. Nagy, L., Tontonoz, P., Alvarez, J. G., Chen, H. & Evans, R. M. Oxidized LDL regulates macrophage gene expression through activation of PPAR γ . *Cell* **93**, 229–240 (1998).
8. Tontonoz, P. *et al.* Terminal differentiation of human liposarcoma cells induced by ligands for peroxisome proliferator-activated receptor γ and the retinoid X receptor. *Proc. Natl Acad. Sci. USA* **94**, 237–241 (1997).
9. Yu, K. *et al.* Differential activation of peroxisome proliferator-activated receptors by eicosanoids. *J. Biol. Chem.* **270**, 23975–23983 (1995).
10. Kliewer, S. A. *et al.* A prostaglandin J2 metabolite binds peroxisome proliferator-activated receptor γ and promotes adipocyte differentiation. *Cell* **83**, 813–819 (1995).
11. Forman, B. M. *et al.* 15-deoxy, delta 12,14-prostaglandin J2 is a ligand for the adipocyte determination factor PPAR γ . *Cell* **83**, 803–812 (1995).
12. Lehman, J. M. *et al.* An antidiabetic thiazolidinedione is a high affinity ligand for the nuclear peroxisome proliferator-activated receptor γ (PPAR γ). *J. Biol. Chem.* **270**, 12953–12956 (1995).

13. Lehmann, J. M., Lenhard, J. M., Oliver, B. B., Ringold, G. M. & Kliewer, S. A. Peroxisome proliferator-activated receptors α and γ are activated by indomethacin and other non-steroidal anti-inflammatory drugs. *J. Biol. Chem.* **272**, 3406–3410 (1997).
14. Mangelsdorf, D. J. *et al.* The nuclear receptor superfamily: the second decade. *Cell* **83**, 835–839 (1995).
15. Wurtz, J. M. *et al.* A canonical structure for the ligand binding domain of nuclear receptors. *Nature Struct. Biol.* **3**, 87–94 (1996).
16. Camp, H. S. & Tafuri, S. R. Regulation of peroxisome proliferator-activated receptor γ activity by mitogen-activated protein kinase. *J. Biol. Chem.* **272**, 13452–13457 (1997).
17. Adams, M., Reginato, M. J., Shao, D., Lazar, M. A. & Chatterjee, V. K. Transcriptional activation by PPAR γ is inhibited by phosphorylation at a consensus mitogen-activated protein kinase site. *J. Biol. Chem.* **272**, 5128–5132 (1997).
18. Hu, E., Kim, J. B., Sarraf, P. & Spiegelman, B. M. Inhibition of adipogenesis through MAP kinase-mediated phosphorylation of PPAR γ . *Science* **274**, 2100–2103 (1992).
19. Allan, G. F. *et al.* Hormone and antihormone induce distinct conformational changes which are central to steroid receptor activation. *J. Biol. Chem.* **267**, 19513–19520 (1992).
20. Berger, J. *et al.* Thiazolidinediones produce a conformational change in peroxisome proliferator-activated receptor- γ : binding and activation correlate with antidiabetic actions in *db/db* mice. *Endocrinology* **137**, 4189–4195 (1996).
21. Lavinsky, R. M. *et al.* Diverse signalling pathways modulate nuclear receptor recruitment of N-CoR and SMRT complexes. *Proc. Natl Acad. Sci. USA* **95**, 2920–2925 (1998).
22. Refetoff, S., Weiss, R. E. & Usala, S. J. The syndromes of resistance to thyroid hormone. *Endocrine Rev.* **14**, 348–399 (1993).
23. Reginato, M. J., Krakow, S. L., Bailey, S. T. & Lazar, M. A. Prostaglandins promote and block adipogenesis through opposing effects on peroxisome proliferator-activated receptor γ . *J. Biol. Chem.* **273**, 1855–1858 (1998).
24. McBroom, L. D. B., Flock, G. & Giguere, V. The nonconserved hinge region and distinct amino-terminal domains of the ROR α orphan nuclear receptor isoforms are required for proper DNA bending and ROR α -DNA interactions. *Mol. Cell. Biol.* **15**, 796–808 (1995).
25. Tzagarakis-Foster, C. & Privalsky, M. L. Phosphorylation of thyroid hormone receptors by protein kinase A regulates DNA recognition by specific inhibition of receptor monomer binding. *J. Biol. Chem.* **273**, 10926–10932 (1998).
26. Bunone, G., Briand, P. A., Miksicek, R. J. & Picard, D. Activation of the unliganded estrogen receptor by EGF involves the MAP kinase pathway and direct phosphorylation. *EMBO J.* **15**, 2174–2183 (1996).
27. Kato, S. *et al.* Activation of the estrogen receptor through phosphorylation by mitogen-activated protein kinase. *Science* **270**, 1491–1494 (1995).
28. McInerney, E. M., Tsai, M. J., O'Malley, B. W. & Katzenellenbogen, B. S. Analysis of estrogen receptor transcriptional enhancement by steroid hormone receptor coactivator. *Proc. Natl Acad. Sci. USA* **93**, 10069–10073 (1996).
29. Schwarz, E. J., Reginato, M. J., Shao, D., Krakow, S. L. & Lazar, M. A. Retinoic acid blocks adipogenesis by inhibiting C/EBP β -mediated transcription. *Mol. Cell. Biol.* **17**, 1552–1561 (1997).
30. Krey, G. *et al.* Fatty acids, eicosanoids, and hypolipidemic agents identified as ligands of peroxisome proliferator-activated receptors by coactivator-dependent receptor ligand assay. *Mol. Endocrinol.* **11**, 779–791 (1997).

Acknowledgements. This work was supported by grants from the National Institute of Diabetes, Digestive, and Kidney Disease at the NIH.

Correspondence and requests for materials should be addressed to M.A.L. (e-mail: lazar@mail.med.upenn.edu).

Swing of the lever arm of a myosin motor at the isomerization and phosphate-release steps

Yoshikazu Suzuki*, Takuo Yasunaga†, Reiko Ohkura*, Takeyuki Wakabayashi† & Kazuo Sutoh*

* Department of Life Sciences, School of Arts and Sciences, University of Tokyo, Komaba 3-8-1, Tokyo 153, Japan

† Department of Physics, School of Science, University of Tokyo, Hongo 7-3-1, Tokyo 113, Japan

In muscle, the myosin head ('crossbridge') performs the 'working stroke', in which ATP is hydrolysed to generate the sliding of actin and myosin filaments. The myosin head consists of a globular motor domain and a long lever-arm domain. The 'lever-arm hypothesis'^{1–5} predicts that during the working stroke, the lever-arm domain tilts against the motor domain, which is bound to actin in a fixed orientation. To detect this working stroke in operation, we constructed fusion proteins by connecting *Aequorea victoria* green fluorescent protein and blue fluorescent protein^{6–8} to the amino and carboxyl termini of the motor domain of myosin II of *Dictyostelium discoideum*, a soil amoeba, and measured the fluorescence resonance energy transfer between the two fluorescent proteins. We show here that the carboxy-terminal fluorophore swings at the isomerization step of the ATP hydrolysis cycle, and then swings back at the subsequent step in which inorganic phosphate is released, thereby mimicking the swing of the lever arm. The

swing at the phosphate-release step may correspond to the working stroke, and the swing at the isomerization step to the recovery stroke.

We replaced the lever-arm domain of the myosin head with either blue or green fluorescent protein (BFP or GFP) by fusing the fluorescent protein to the C terminus of the motor domain (residues 1–761) of *Dictyostelium* myosin II (ref. 4), designated as S1dC (ref. 10), through a spacer consisting of three glycine residues. We fused the second fluorescent protein to the N terminus of S1dC through the spacer, and thereby produced a chimaeric S1dC, GS1dCB or BS1dCG (Fig. 1). The C-terminal fluorescent protein in the chimaeric S1dC is expected to mimic motion of the lever arm, which is normally connected to the C-terminal domain (residues 692–761) of S1dC. Fluorescence resonance energy transfer (FRET) within the chimaeras could therefore be used to detect the swing motion which is predicted to occur from structural studies^{11–13}.

Excitation of BFP in GS1dCB with light of 360 nm wavelength resulted in blue emission at 430–470 nm, whereas FRET between BFP and GFP resulted in a strong green emission from GFP at 490–540 nm. On addition of ATP, the blue emission from BFP increased and the green emission from GFP decreased (Fig. 2a). The calculated distance between GFP and BFP of GS1dCB increased from 3.6 to 5.1 nm when the chimaera bound ATP (Table 1). Upon addition of ADP to GS1dCB, however, the calculated distance changed only slightly (Fig. 2b and Table 1). FRET measurements obtained from the other type of chimaera, BS1dCG, were very similar. Binding of nucleotides conjugated to the fluorophore Cy3, Cy3-ATP or Cy3-ADP, to the ATPase site of GS1dCB or BS1dCG resulted in energy transfer from GFP to the Cy3-nucleotide upon excitation of GFP at 480 nm. The calculated distances from Cy3-ATP (or Cy3-ADP) in the ATPase site to the fluorophores at the N and C termini of S1dC were 5.3 and 5.7 nm, respectively. Those from Cy3-ADP in the ATPase site to the fluorophores at the N and C termini were 5.7 and 6.2 nm, respectively.

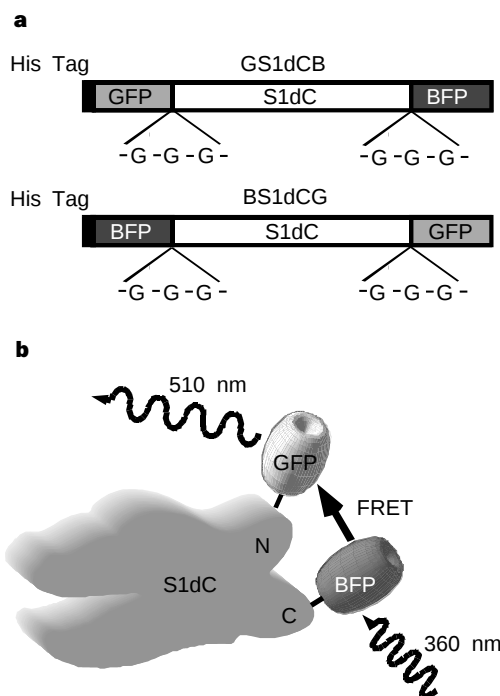


Figure 1 Diagrams of the domain of chimaeras made from the incorporation of fluorescent proteins into S1dC. **a**, The GFP, BFP and S1dC domains were connected by flexible spacers consisting of three glycine residues to generate GS1dCB and BS1dCG. **b**, When BFP was excited by light at 360 nm, intramolecular fluorescence resonance energy transfer (FRET) took place and green light was emitted at 510 nm from GFP. The intensity of the emission from GFP depended on the distance between GFP and BFP in the chimaeras.

In the cyclical hydrolysis of ATP by myosin, several intermediate states defined by their characteristic optical properties have been identified^{14,15}. Upon binding ATP, the chimaeric S1dC, initially in the M state (without any bound nucleotide), passes through the M.ATP and M*.ATP states quickly, and then hydrolyses ATP to assume the most stable intermediate state, the M**.ADP.P_i state (asterisks indicate different conformational states of myosin; see Fig. 2d legend). It releases P_i slowly to assume the M.ADP state, and then releases ADP quickly to resume the M state¹⁵ (Fig. 2d). Thus, according to our FRET measurements, the distance between GFP and BFP increased at a step between the M and M**.ADP.P_i states. This increase probably arose from the swing of the BFP fluorophore which mimicked the motion of the lever-arm domain, as discussed below. This swing of the fluorophore was reversed at the subsequent P_i-releasing step between the M**.ADP.P_i and M.ADP states (Fig. 2d). The fluorophore did not swing further at the ADP-releasing step between the M.ADP and M states. Because the working stroke is expected to occur at the P_i-releasing step¹⁴, the observed reverse swing seems to correspond to this stroke, whereas the swing at a step between the M and M**.ADP.P_i states seems to correspond to the recovery stroke that induces the relaxed state, ready for the subsequent working stroke.

We further investigated the swing step between the M and M**.ADP.P_i states by using G457A and E459A mutants of GS1dCB. Biochemical studies of the G457A and E459A mutations of *Dictyostelium* myosin¹⁶ indicate that the addition of ATP traps the G457A mutant in a stable state somewhere between the M.ATP and M*.ATP states, whereas the E459A mutant irreversibly binds ATP and is frozen in the M*.ATP state (Fig. 2d). The FRET efficiency of the E459A and G457A mutants did not change upon addition of ATP, and remained comparable to that in the wild-type chimaera with and without ATP, respectively (Fig. 2c), indicating that the G457A and E459A mutants were trapped in the structural states before and after the swing, respectively. Thus, the swing step corresponds to the isomerization step between the M.ATP and M*.ATP states (Fig. 2d).

We monitored the time course of the swing and reverse swing by

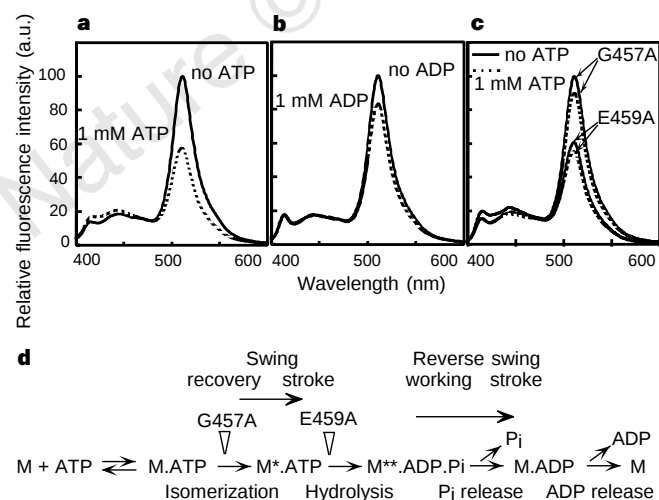


Figure 2 The relationship of FRET emission spectra in chimaeric S1dC to the hydrolysis of ATP. **a**, Emission spectra of GS1dCB excited at 360 nm in the presence and absence of ATP. **b**, Emission spectra of GS1dCB in the presence and absence of ADP. **c**, Emission spectra of G457A and E459A mutants of GS1dCB in the presence and absence of ATP. The spectra of the G457A mutant were similar to that of the wild-type chimaera without ATP, whereas those of the E459A mutant were similar to that of the wild-type with ATP. **d**, Intermediate states in the ATPase cycle of myosin^{14,15}. M, M* and M** represent myosin in different conformational states as defined by their optical properties. Steps blocked by the G457A and E459A mutations are indicated by open arrowheads. The swing and reverse swing occur at the stages indicated.

stopped-flow measurements. Following the addition of ATP to GS1dCB, the FRET-induced emission from GFP rapidly decreased. After ATP hydrolysis, the emission from GFP gradually returned to its initial level (Fig. 3a). The time course of the initial rapid decrease in emission (Fig. 3b) was well fitted by a curve with an exponential component and a slow linear component. The slow linear component may have arisen from photobleaching of fluorophores as it was observed even without ATP. The rate constant derived from the exponential curve was linearly dependent on the concentration of ATP, with an apparent second-order rate constant of $0.44 \mu\text{M}^{-1} \text{s}^{-1}$. At higher ATP concentrations, the rate constant reached a plateau towards a maximum of 33s^{-1} (Fig. 3c). This is the first-order rate constant for the swing of the recovery stroke. These rate constants are close to those for isomerization of S1dC induced by ATP binding ($0.17 \mu\text{M}^{-1} \text{s}^{-1}$ and 17s^{-1})¹⁷. The first-order rate constant of the reverse swing, calculated from the recovery phase of the sensitized emission, was 0.022s^{-1} , which agrees well with the steady state rate of P_i-release by GS1dCB (0.029s^{-1}).

The crystal structure of S1dC has two distinct states. When complexed with magnesium salts of ADP, ATP-γS, AMPPNP or of ADP.BeF₃, it assumes an 'ADP' structure. When complexed with magnesium derivatives of ADP.V_i or of ADP.AlF₄, it assumes a 'vanadate', or V_i structure^{11–13}. To correlate these two types of crystal structure with the two structural states before and after the swing, we carried out FRET measurements in the presence of various nucleotides and their analogues. The calculated distance of the fluorophores of GS1dCB complexed with Mg.ADP.AlF₄ or Mg.ADP.V_i was 5.3–5.4 nm (Table 1). The results show that the orientation of BFP and GFP in these complexes is very similar to that in the M**.ADP.P_i state, and indicate that GS1dCB in the M**.ADP.P_i state assumes the V_i structure. This is consistent with the results of previous biochemical^{18,19} and X-ray studies^{12,13}. The calculated distance was 3.8–3.9 nm for the complex with Mg.ATPγS or Mg.ADP (Table 1), consistent with X-ray results showing that both complexes assume the ADP structure¹¹. For the complex with Mg.ADP.BeF₃ or Mg.AMPPNP, however, there was an apparent discrepancy between the FRET measurements and the crystal structures^{11,13}. The calculated distance (4.3–4.4 nm) was intermediate between the values expected for the ADP and V_i structures, suggesting that the two types of structural states may co-exist, or that the chimaeric S1dC can assume a third structural state. It is likely that the M**.ADP.P_i state corresponds to the V_i structure, whereas the M or M.ADP state corresponds to the ADP structure.

Using crystal structures of GFP²⁰ and S1dC^{11,13}, we modelled the locations of GFP and BFP in GS1dCB (assuming the M**.ADP.P_i or M.ADP state) that satisfy the calculated distances between the fluorophores and the ATPase site (see Methods). The calculated locations of the fluorophores were superimposed on the reconstructed myosin head, assuming either the V_i structure or the ADP structure (Fig. 4). The BFP fluorophores attached to the C terminus of S1dC in the M**.ADP.P_i state (shown in pink in Fig. 4) and in the M.ADP state (orange) overlapped with the lever-arm domains in the V_i structure (coloured dark orange) and in the ADP structure (green), respectively. Thus, the observed swing of the BFP fluorophore of GS1dCB at the isomerization step (corresponding to the

Table 1 Distances between BFP and GFP in GS1dCB

Complex	FRET efficiency	Distance (nm)*	Distance change (nm)
No nucleotide	0.439	3.6	–
ATP (1 mM)	0.082	5.1	+1.5
ADP (1 mM)	0.333	3.8	+0.2
Mg.ADP/BeF ₃	0.172	4.4	+0.8
Mg.ADP/V _i	0.085	5.3	+1.7
Mg.ADP/AlF ₄	0.062	5.4	+1.8
Mg.AMPPNP (1 mM)	0.209	4.3	+0.7
Mg.ATPγS (1 mM)	0.301	3.9	+0.3

* Standard deviations were less than 0.14 nm.

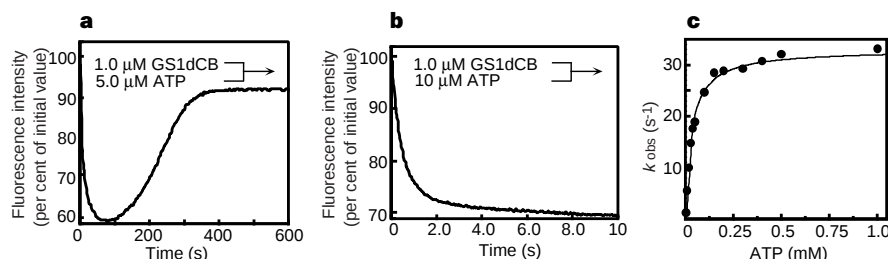


Figure 3 Time course of changes in the FRET-induced emission from GFP. **a**, Rapid decrease and slow recovery of the GFP emission upon rapid mixing of 1 μ M GS1dCB with an equal volume of 5 μ M ATP (concentrations before mixing). **b**, The time course of the rapid phase at 10 μ M ATP is best fitted by a curve with a

single exponential component and a slow linear component. From the single exponential curve, an apparent first-order rate constant of the rapid decrease in the emission (k_{obs}) was obtained. **c**, Plot of rate constants (k_{obs}) against ATP concentration.

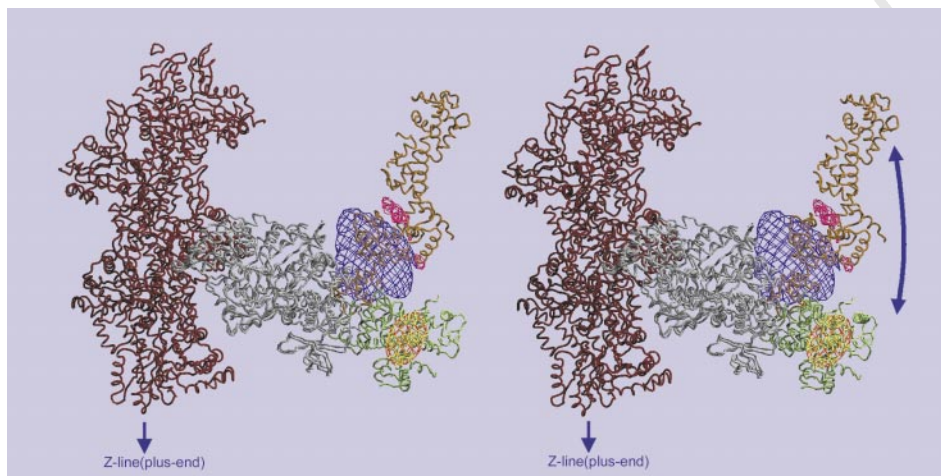


Figure 4 Stereo views of a complex of an actin filament and a myosin head, prepared by using program O (ref. 26). The 'meshed' structures represent the probable positions of the fluorophores. According to the calculation for the probability distribution of fluorophores, the probability that they should be inside these mesh structures was 80%. Blue mesh: the position of the fluorophore fused to the N terminus of S1dC. Pink/orange mesh: the position of the fluorophore fused to the C terminus of S1dC—pink in the M^{ADP} state and orange in

the M^{ADP} state. Actin is shown in dark brown. S1dC is shown in grey, irrespective of the bound nucleotide. The lever-arm domain in the V_i structure is in dark orange, and is green in the ADP structure. The fluorophore positions shown in pink and orange fit well with the lever-arm domains in the corresponding structures. The blue double-headed arrow indicates putative swings of the lever-arm. The plus-end of the actin filament is located downwards, so myosin moves downwards during force generation.

recovery stroke) is consistent with the motion of lever-arm domain from the ADP structure to the V_i structure, whereas the reverse swing at the P_i-releasing step (corresponding to the working stroke) is consistent with motion in the reverse direction. This latter motion of the lever-arm domain can generate sliding of myosin along actin filaments in the appropriate direction for muscle contraction (towards the plus-end, Z-line direction). □

Methods

Sample preparation and steady-state fluorescence measurements. Red-shifted GFP was constructed by introducing F64L/S65T mutations in pGFP (Clontech). BFP was cloned from the pEBFP (Clontech) by polymerase chain reaction (PCR). The red-shifted GFP and BFP were fused to S1dC, and the recombinant proteins were expressed in *Dictyostelium* cells and purified as described^{10,16,21}. Steady-state fluorescence was measured with a Perkin-Elmer LS50B fluorophotometer in a solution comprising 20 mM MOPS, pH 7.4, 50 mM KCl and 2 mM MgCl₂ at 20 °C. Protein concentrations were 0.1–0.2 μ M. Fluorophores were excited at wavelengths of 360 nm (BFP) and 480 nm (GFP). FRET efficiencies (E) were determined from the FRET-induced emission of GFP according to²²: $E = \{(\text{ratio})^A - \epsilon^A(360)/\epsilon^A(480)\} \times \epsilon^A(480)/\epsilon^D(360)$, where the superscripts D and A refer to donor (BFP) and acceptor (GFP), $\epsilon^A(360)$, $\epsilon^A(480)$ and $\epsilon^D(360)$ are the molar extinction coefficients of the acceptor and donor at the designated wavelength. The emission ratio of acceptor is defined as $(\text{ratio})^A = F_{\text{FRET}}^A/F_{\text{DIR}}^A$, where F_{FRET}^A is the fluorescence intensity of the acceptor (GFP) at 510 nm when the donor (BFP) is

excited, and F_{DIR}^A is the fluorescence intensity of acceptor (GFP) at 510 nm when it is directly excited at 480 nm. Distances (R) between fluorophores of BFP and GFP were calculated using Forster's equation²³ ($R = R_0(E^{-1} - 1)^{1/6}$) assuming that the rotational mobility of each probe did not change when a nucleotide was added. R_0 is Forster's critical distance at which the transfer efficiency is 50% and was obtained (in nm) from the equation: $R_0^6 = (8.79 \times 10^{17})n^{-4}k^2Q_DJ$, where J is the numerical value of the overlap integral between BFP and GFP when expressed in M⁻¹ cm³, which was 7.2×10^{-14} . Quantum yield (Q_D) was 0.15 for BFP (0.55 for quinine sulphate as standard). The refractive index (n) and orientation factor (k^2) were assumed to be 1.4 and 2/3, respectively. The calculated R_0 (between GFP and BFP) was 3.4 nm. Distances between GFP fluorophores at the N and C termini of S1dC and Cy3-nucleotide in the ATPase site were calculated from FRET efficiencies between GFP and Cy3 detected at 510 nm by excitation at 480 nm. The J value between GFP and Cy3 was calculated as 4.4×10^{-13} M⁻¹ cm³. Q_D (GFP) was determined as 0.47 (0.75 for fluorescein as a standard). The calculated R_0 (between GFP and Cy3) was 5.6 nm. In this case, FRET efficiencies were determined by measuring the fluorescence intensity of the donor (GFP) in the presence (F_{DA}) and absence (F_{D}) of the acceptor (Cy3-nucleotide) ($E = 1 - F_{\text{DA}}/F_{\text{D}}$). The inner-filter effect of Cy3 was corrected as described²⁴. Conditions for FRET measurements were the same as above. Chimaeric S1dC was complexed with MgADP.BeFx, MgADP.V_i or MgADP.AIF₄, as described^{18,19}.

Fast kinetics. We carried out kinetic measurements with a stopped-flow spectrometer (Applied Photophysics SX18); the excitation wavelength was 360 nm. Emission from GFP was detected through a 570-nm high-pass filter to

eliminate emission from BFP. Conditions were the same as for the steady-state measurements apart from the protein concentration (1.0 μM).

Model building. We constructed structural models of GS1dCB in the $M^{**}\text{ADP.P}_i$ and $M\text{ADP}$ states using coordinates of the V_i and ADP structures of S1dC^{11–13} and those of GFP²⁰. As many residues of the C-terminal domain of S1dC are missing in the S1dC data, they were restored by fitting with the chicken S1 structure using the program Eos²⁵. The myosin head with the lever-arm domain, assuming the V_i or ADP structure, was then reconstructed from the V_i or ADP structure of the *Dictyostelium* S1dC^{11–13} and the structure of the lever-arm domain of chicken S1 (ref. 3) as before⁵. Next, we calculated the probability distribution of the position of fluorophore of BFP or GFP. The calculation was based on FRET data and the following assumptions: (1) the chain of GFP or BFP cannot occupy the same position as the S1dC chain; (2) the distance between the fluorophore of GFP and the N terminus of S1dC is 3.5 nm, and that between the fluorophore of BFP and the C terminus of S1dC is 2.5 nm, consistent with the atomic structure of GFP²⁰; (3) each of the distances between the points (GFP fluorophore and N terminus of S1dC, GFP fluorophore and Cy3, BFP fluorophore and C terminus of S1dC, BFP fluorophore and Cy3, BFP fluorophore and GFP fluorophore) can vary around the measured value as though the points are connected by springs for which the spring constant is $k_B T / (0.25 \text{ nm})^2$. By introducing this length variability, the error of the distance obtained by FRET measurements (the range of the standard deviation was 0.02–0.14 nm) and the error in the initial estimate of the position of Cy3 can be accommodated. Then we calculated the probability distribution of the fluorophore using two types of S1dC structure found in the PDB (Brookhaven Protein Data Bank) files under accession numbers 1mmd (S1dC complexed with MgADP.BeF_x) and 1vom (S1dC complexed with MgADP.V_i). The best probability distribution was obtained when we used the PDB files 1vom and 1mmd for the $M^{**}\text{ADP.P}_i$ and $M\text{ADP}$ states, respectively.

Received 27 May; accepted 10 September 1998.

- Cooke, R. The mechanism of muscle contraction. *CRC Crit. Rev. Biochem.* **21**, 53–118 (1986).
- Rayment, I. *et al.* Structure of the actin-myosin complex and its implications for muscle contraction. *Science* **261**, 58–65 (1993).
- Rayment, I. *et al.* Three-dimensional structure of myosin subfragment-1: a molecular motor. *Science* **261**, 50–58 (1993).
- Spudich, J. A. How molecular motors work. *Nature* **372**, 515–518 (1994).
- Holmes, K. C. The swinging lever-arm hypothesis of muscle contraction. *Curr. Biol.* **7**, 112–118 (1997).
- Prasher, D. C., Eckenrode, V. K., Ward, W. W., Prendergast, F. G. & Cormier, M. J. Primary structure of the *Aequorea victoria* green-fluorescent protein. *Gene* **111**, 229–233 (1992).
- Heim, R. & Tsien, R. Y. Engineering green fluorescent protein for improved brightness, longer wavelengths and fluorescence resonance energy transfer. *Curr. Biol.* **6**, 178–182 (1996).
- Miyawaki, A. *et al.* Fluorescent indicators for Ca^{2+} based on green fluorescent proteins and calmodulin. *Nature* **388**, 882–887 (1997).
- Warrick, H. M., DeLozanne, A., Leinwand, L. A. & Spudich, J. A. Conserved protein domains in a myosin heavy chain gene from *Dictyostelium discoideum*. *Proc. Natl Acad. Sci. USA* **83**, 9433–9437 (1986).
- Itakura, S. *et al.* Force-generating domain of myosin motor. *Biochem. Biophys. Res. Commun.* **196**, 1504–1510 (1993).
- Gulick, A. M., Bauer, C. B., Thoden, J. B. & Rayment, I. X-ray structures of the MgADP , $\text{MgATP}\gamma\text{S}$, and MgAMPPNP complexes of the *Dictyostelium discoideum* myosin motor domain. *Biochemistry* **36**, 11619–11628 (1997).
- Smith, C. A. & Rayment, I. X-ray structure of the magnesium(II)-ADP:vanadate complex of the *Dictyostelium discoideum* myosin motor domain to 1.2 Å resolution. *Biochemistry* **35**, 5404–5417 (1996).
- Fisher, A. J. *et al.* X-ray structures of the myosin motor domain of *Dictyostelium discoideum* complexed with MgADP.BeF_x and MgADP.AIF_x . *Biochemistry* **34**, 8960–8972 (1995).
- Lynn, R. W. & Taylor, E. W. Mechanism of adenosine triphosphate hydrolysis of actomyosin. *Biochemistry* **10**, 4617–4624 (1971).
- Bagshaw, C. R. & Trentham, D. R. The characterization of myosin-product complexes and of product-release steps during the magnesium ion-dependent adenosine triphosphatase reaction. *Biochem. J.* **141**, 331–349 (1974).
- Sasaki, N., Shimada, T. & Sutoh, K. Mutational analysis of the switch II loop of *Dictyostelium* myosin II. *J. Biol. Chem.* **273**, 20334–20340 (1998).
- Kuhlman, P. A. & Bagshaw, C. R. ATPase kinetics of the *Dictyostelium discoideum* myosin II motor domain. *J. Muscle. Res. Cell. Motil.* **19**, 491–504 (1998).
- Werber, M. M., Peyser, Y. M. & Muhrad, A. Characterization of stable beryllium fluoride, aluminum fluoride, and vanadate containing myosin subfragment 1-nucleotide complexes. *Biochemistry* **31**, 7190–7197 (1992).
- Maruta, S., Henry, G. D., Sykes, B. D. & Ikebe, M. Formation of the stable myosin-ADP-aluminum fluoride and myosin-ADP-beryllium fluoride complexes and their analysis using ^{19}F NMR. *J. Biol. Chem.* **268**, 7093–7100 (1993).
- Ormo, M. *et al.* Crystal structure of the *Aequorea victoria* green fluorescent protein. *Science* **273**, 1392–1395 (1996).
- Shimada, T., Sasaki, N., Ohkura, R. & Sutoh, K. Alanine scanning mutagenesis of the switch I region in the ATPase site of *Dictyostelium discoideum* myosin II. *Biochemistry* **36**, 14037–14043 (1997).
- Clegg, R. M. Fluorescence resonance energy transfer and nucleic acids. *Methods Enzymol.* **211**, 353–388 (1992).
- Stryer, L. Fluorescence energy transfer as a spectroscopic ruler. *Annu. Rev. Biochem.* **47**, 819–846 (1978).

- Miki, M. Detection of conformational changes in actin by fluorescence resonance energy transfer between tyrosine-69 and cysteine-374. *Biochemistry* **30**, 10878–10884 (1991).
- Yasunaga, T. & Wakabayashi, T. Extensible and object-oriented system Eos supplies a new environment for image analysis of electron micrographs of macromolecules. *J. Struct. Biol.* **116**, 155–160 (1996).
- Jones, T. A., Zou, J. Y., Cowan, S. W. & Kjeldgaard, M. Improved methods for building protein models in electron density maps and the location of errors in these models. *Acta Crystallogr. A* **47**, 110–119 (1991).

Acknowledgements. We thank K. Oiwa for Cy3 nucleotides, M. Miki for discussions about the FRET experiments, and M. Steward, C. Bagshaw and P. Kuhlman for critically reading the manuscript.

Correspondence and requests for materials should be addressed to K.S. (e-mail: kksutoh@chem.t.u.tokyo.ac.jp).

Elastic bending and active tilting of myosin heads during muscle contraction

Ian Dobbie*, Marco Linari†, Gabriella Piazzesi‡, Massimo Reconditi†, Natalia Koubassova‡, Michael A. Ferenczi§, Vincenzo Lombardi† & Malcolm Irving*

* Randall Institute, King's College London, London WC2B 5RL, UK

† Dipartimento di Scienze Fisiologiche, Università degli Studi di Firenze, Viale GB Morgagni 63, I-50134 Firenze, Italy

‡ Institute of Mechanics, Moscow State University, Moscow 119899, Russia

§ National Institute for Medical Research, Mill Hill, London NW7 1AA, UK

Muscle contraction is driven by a change in shape of the myosin head region that links the actin and myosin filaments^{1,2}. Tilting of the light-chain domain of the head with respect to its actin-bound catalytic domain is thought to be coupled to the ATPase cycle^{3–6}. Here, using X-ray diffraction and mechanical data from isolated muscle fibres, we characterize an elastic bending of the heads that is independent of the presence of ATP. Together, the tilting and bending motions can explain force generation in isometric muscle, when filament sliding is prevented. The elastic strain in the head is 2.0–2.7 nm under these conditions, contributing 40–50% of the compliance of the muscle sarcomere. We present an atomic model for changes in head conformation that accurately reproduces the changes in the X-ray diffraction pattern seen when rapid length changes are applied to muscle fibres both in active contraction and in the absence of ATP. The model predictions are relatively independent of which parts of the head are assumed to bend or tilt, but depend critically on the measured values of filament sliding and elastic strain.

We measured the elastic properties of the actin and myosin filaments and myosin heads in actively contracting single muscle fibres by imposing a sinusoidal length change of about 5 nm per half-sarcomere, peak-to-peak with period 320 μs (Fig. 1). The force oscillated between about 0.6 and 1.6 T_0 , where T_0 is the steady force during isometric contraction (Fig. 1a). No phase difference could be detected between the force and sarcomere length changes with the 8- μs sampling interval (Fig. 1b). Thus the length changes are sufficiently fast, and the compliance of the fibre attachments sufficiently small, that the force response is purely elastic⁷; neither the rapid force recovery associated with the working stroke in the attached myosin head⁸, nor the slower detachment of myosin heads from actin and reattachment to it^{9,10} contribute to the force response during the 320- μs cycle.

Elastic deformation of myosin heads alters the intensity of the M3 X-ray reflection (I_{M3}) arising from the 14.5-nm spacing between adjacent sets of myosin heads along the filaments^{11,12}. I_{M3} decreased by $25 \pm 3\%$ (mean $\pm$ s.e., 5 fibres) during the stretch phase of the length-change cycle (Fig. 1c, filled circles). The change in I_{M3} , recorded at 20- μs time resolution, was exactly 180° out of phase

with the length and force signals. Thus the motion of the myosin heads in this protocol is purely elastic, like the mechanical response.

The compliance of the myosin filaments was estimated from the spacing of the M3 reflection, and that of the actin filaments from the spacing of the 5.9-nm layer line reflection from the actin helix, by comparing 100- μ s time-frames at the peak and trough of the force oscillation. The strains in the myosin and actin filaments were $0.12 \pm 0.01\%/T_0$ and $0.30 \pm 0.13\%/T_0$ respectively ($\pm$ s.e., 5 fibres), in agreement with the spacing changes of a wider range of X-ray reflections during slow stretch of whole muscles^{13,14}. Our results show that these spacing changes are not caused by attachment or detachment of the myosin heads or by the working stroke in the

attached heads, processes that are slower than the length-change cycle used here.

The total compliance of the muscle fibre segment illuminated by the X-ray beam during the 320- μ s length-change cycle was $5.1 (\pm 0.3) \text{ nm}/T_0$ in each half-sarcomere. This includes the contributions of myosin and actin filaments and the myosin heads that crosslink them in the zone of filament overlap. The contributions of the myosin and actin filaments were calculated from the X-ray measurements of average filament strain, taking into account the strain distribution along the filaments^{15,16}, as 0.8 and 2.3 nm/T_0 respectively. Neglecting possible small contributions from other sarcomeric structures, the remaining 2.0 nm/T_0 of the half-sarcomere compliance in active contraction should be due to the myosin heads. The main source of uncertainty in this estimate is our value of $0.30 \pm 0.13\%/T_0$ for the average strain in the actin filament. If this were replaced by the more precise published values, either $0.23 \pm 0.01\%/T_0$ from the spacing change of the 2.7 nm actin-based X-ray reflection during slow stretch of whole muscles¹³, or $0.21 \pm 0.03\%/T_0$ from the variation of sarcomere compliance with sarcomere length in isolated fibres¹⁶, the myosin head compliance would be about 2.7 nm/T_0 .

Elastic deformation of myosin heads should still be present in the absence of ATP (in rigor), when all the myosin heads are bound to actin^{17,18}. To test this prediction, we applied a peak-to-peak length change of 2.4 nm per half-sarcomere and period of 320 μ s to fibres that had been permeabilized and depleted of ATP (Fig. 2). The

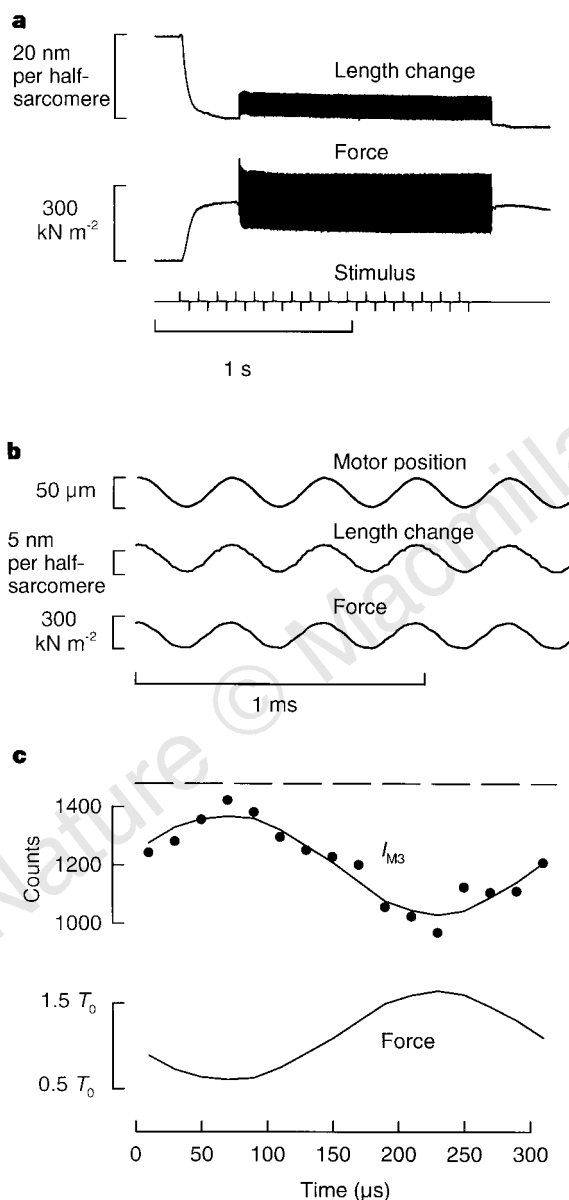


Figure 1 Changes in force and intensity of the M3 X-ray reflection (I_{M3}) produced by rapid length oscillations. **a**, Length change (nm per half-sarcomere) of the segment in the X-ray beam and force, when 4,000 oscillations with period of 320 μ s were imposed during electrical stimulation of an intact muscle fibre; cross-sectional area, 20,100 μm^2 ; length, 6.63 mm; mean sarcomere length in the 2.34 mm segment, 2.09 μm . **b**, Changes in motor position, segment length and force, sampled at 8- μ s intervals 1 s after the start of the oscillations. **c**, I_{M3} (filled circles) and force (continuous lines; that in the upper panel inverted and scaled to fit I_{M3}) averaged over the 4,000 cycles from 61 tetani in 5 fibres; peak-to-peak length change $5.27 \pm 0.12 \text{ nm}$ per half-sarcomere. Dashed line: I_{M3} without applied length changes in the same fibres (11 tetani).

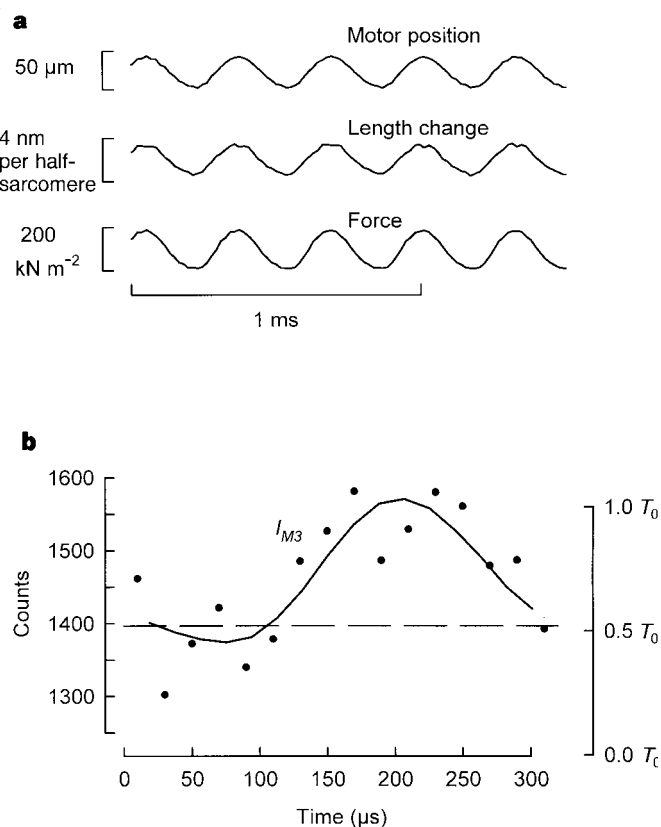


Figure 2 Changes in force and I_{M3} produced by length oscillations with 320 μ s period in rigor. **a**, Motor position, length change of segment in the X-ray beam, and force sampled at 16- μ s intervals starting 30 ms after the start of 4,000 stretch-shortening cycles. Before the oscillations force was $0.51 T_0$; fibre length, 3.52 mm; mean sarcomere length in the 2.16 mm segment, 2.15 μm . Cross-sectional area: 21,100 μm^2 in the intact fibre, 24,800 μm^2 in rigor. **b**, I_{M3} (filled circles) and force, summed over the 4,000 cycles, from 200 runs in 8 fibres, peak-to-peak length change, $2.4 \pm 0.1 \text{ nm}$ per half-sarcomere. Dashed line: I_{M3} without imposed length changes in the same fibres (29 runs); force was similar to the minimum value during oscillations.

resulting force modulation was $0.57 \pm 0.02 T_0$ ($\pm$ s.e., 8 fibres), corresponding to a half-sarcomere compliance of $4.3 \pm 0.2 \text{ nm}/T_0$. There was no detectable phase lag between the sarcomere length and force signals (Fig. 2a), as expected for a purely elastic response. I_{M3} increased by $14 \pm 4\%$ during the stretch (Fig. 2b); this change is in the opposite direction to that seen in active contraction (Fig. 1c), as reported previously for slower length changes¹⁹. The spacings of the X-ray reflections could not be measured precisely in rigor but, assuming that filament compliances have the same value as in active contraction²⁰, the compliance of the myosin heads in rigor would be $1.2 \text{ nm}/T_0$, or about $1.9 \text{ nm}/T_0$ if we use published values of actin filament compliance^{13,16} as above.

We constructed an atomic model for the compliance of the myosin head based on elastic bending of its light-chain domain (Fig. 3). Head conformation in an unstrained rigor fibre (Fig. 3a) was assumed to be the same as that deduced from electron micrographs of isolated actin filaments decorated with myosin head fragments³. When a fibre in rigor is stretched, the light-chain domain of the head is displaced towards the M line of the sarcomere (upwards in Fig. 3b), so the axial position of its centre of mass moves closer to that of the catalytic domain. The mass projection of the head onto the filament axis becomes narrower, so I_{M3} increases, as we observed. The size of the increase depends on the myosin head compliance; the observed increase of $14 \pm 4\%$ (Fig. 2b) requires a head compliance of $1.3 \pm 0.4 \text{ nm}/T_0$, in agreement with the value of $1.2\text{--}1.9 \text{ nm}/T_0$ in rigor estimated above.

In active contraction, I_{M3} decreases during a stretch (Fig. 1c). This can only be explained by bending of the light-chain domain if it has already tilted substantially away from the rigor conformation during active contraction. Following current models for the changes

in myosin head conformation during the ATPase cycle^{3–6}, we represented this active motion as tilting of the entire light-chain domain with respect to the actin-bound catalytic domain. Starting from the unstrained rigor conformation, the light-chain domain must be tilted by 22° towards the M line of the half-sarcomere to give the narrowest projection of the mass of the whole head on the filament axis, and thus the maximum value of I_{M3} . Experimentally, I_{M3} is maximized by shortening of $\sim 1 \text{ nm}$ per half-sarcomere from the active isometric conformation²¹, which corresponds to tilting the light-chain domain away from the M line by $\sim 8^\circ$. Thus the total tilt between rigor and active isometric contraction is about 30° .

In the experiment shown in Fig. 1, the average force is higher than in isometric contraction and I_{M3} is lower (Fig. 1a, c). This suggests that the light-chain domain is tilted upwards by a further $\sim 10^\circ$ from its isometric conformation during this protocol (Fig. 3c). When an active fibre is stretched, bending of the light-chain domain displaces its centre of mass away from that of the catalytic domain, so I_{M3} decreases, as we observe. The decrease of $25 \pm 3\%$ in the experiment shown in Fig. 1 requires a myosin head compliance in active contraction of $2.3 \pm 0.3 \text{ nm}/T_0$, in agreement with the value $2.0\text{--}2.7 \text{ nm}/T_0$ estimated above from the filament and sarcomere compliances.

According to this model, the tilting motion associated with the ATPase cycle can still occur during isometric contraction, when filament sliding is prevented and the strain in the filaments is constant. In these conditions, the local conformational change at the junction between the catalytic and light-chain domains is accommodated by elastic bending of the light-chain domain, so that the head-rod junction does not move with respect to the actin-binding site. The resulting change in head conformation would be difficult to detect by low-resolution structural methods.

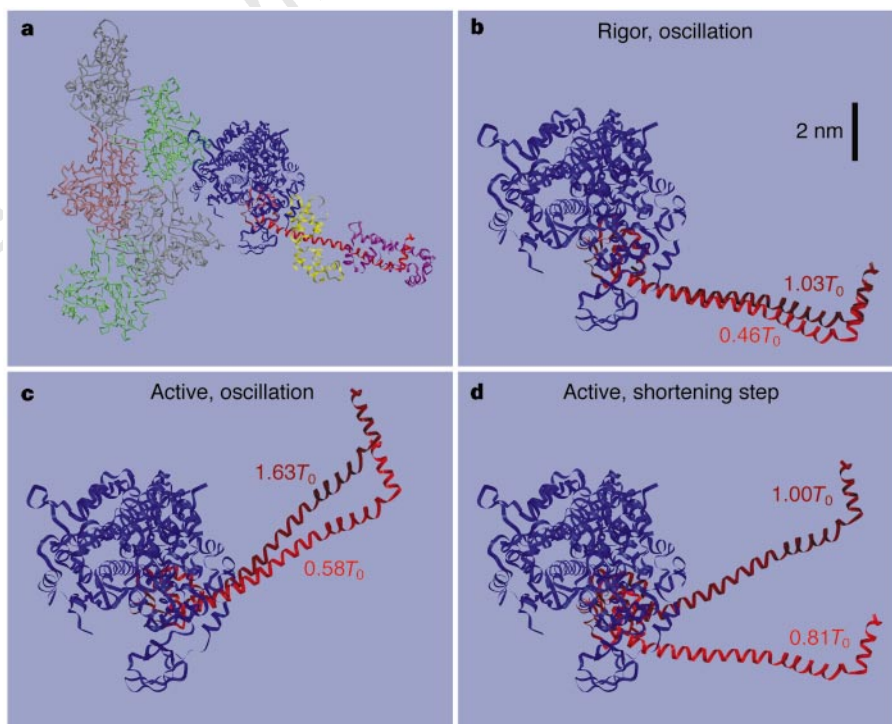


Figure 3 Structural model for elastic bending and ATP-driven tilting of the light-chain domain of the myosin head. **a**, Catalytic (heavy-chain residues 1–707, blue; 707–770, pink) and light-chain domains (heavy-chain residues 771–843, red; essential light chain, yellow; regulatory light chain, magenta) of the myosin head, bound to a vertical actin filament (grey, green, brown) in rigor, with the sarcomeric M line at the top of the figure. In **b–d**, to which the scale bar in **b** applies, actin and light chains are omitted for clarity, catalytic domain (blue) is in the same conformation as **a**, and the moving part of heavy chain (residues 707–843) is shown in red. **b**, Elastic bending of light-chain domain in rigor (Fig. 2), between

force $0.46 T_0$ (light red) and $1.03 T_0$ (dark red). **c**, Elastic bending of light-chain domain during active contraction (Fig. 1) between $0.58 T_0$ (light red) and $1.63 T_0$ (dark red). **d**, Head conformation before (force, $1.00 T_0$, dark red) and 2 ms after (force, $0.81 T_0$; light red) a shortening step of 5.6 nm per half-sarcomere during active contraction. The conformations shown were calculated assuming a myosin head compliance of $1.2 \text{ nm}/T_0$ in rigor and $2.0 \text{ nm}/T_0$ in active contraction. The tilt angle between the light-chain and catalytic domains, relative to that in rigor, was 40° in **c**, 30° at $1.00 T_0$ and 0° at $0.81 T_0$ in **d**. Graphics prepared using Raster3D³⁰. Coordinates kindly provided by I. Rayment³.

Much larger changes in head conformation would occur if filament sliding were allowed. When a rapid shortening step is imposed during active contraction, I_{M3} decreases during the rapid force recovery that follows the step^{12,22}, and we suggested previously that this elementary force-generating process is due to a change in conformation of the myosin heads. When a shortening step of 5.6 nm per half-sarcomere is imposed from the isometric force level (T_0), force recovers to 0.81 T_0 by 2 ms after the step¹², so the change in elastic strain is only 1.0 nm and the remaining 4.6 nm must be accounted for by tilting of the light-chain domain (Fig. 3d). The required tilt angle is 30°, which brings the head to its rigor conformation. Moreover, the observed decrease in I_{M3} , $29 \pm 6\%$ (mean $\pm$ s.e., $n = 7$; see Fig. 4 of ref. 12), gives an independent estimate of the required tilt angle, $28.5 \pm 2.5^\circ$, which is in good agreement with that calculated from the filament displacement.

This structural model for the function of the myosin motor has only two parameters: the tilt of the light-chain domain and the elastic strain. The changes in these two parameters can be calculated from measurements of the filament displacement, the force change, and the compliances of the myosin head and sarcomere. The observed changes in I_{M3} were consistent with the predictions of this structural model in a wide range of conditions (our Figs 1 and 2, and Fig. 4 of ref. 12). I_{M3} is relatively insensitive to the assumptions about which parts of the head bend or tilt. For example, a localized elastic element at the junction between the catalytic and light-chain domains would have an effect very similar to that of uniform bending of the light-chain domain. Further, tilting of the whole head around its actin-binding site would produce a decrease in I_{M3} in the protocol in Fig. 3d, similar to that produced by tilting at the catalytic/light-chain domain boundary, for the same amount of filament sliding. The critical features of the model for reproducing the observed changes in I_{M3} are the value of the myosin head compliance and the overall tilt of the heads towards the M line during isometric contraction.

The elastic deformation described here for myosin is likely to be a fundamental property of motor proteins, including those of the kinesin and dynein families. All these motors are thought to work by an ATP-driven change in conformation of a domain analogous to the myosin head. Many of them have a region analogous to the myosin light-chain domain which could amplify small conformational changes around the ATP-binding site into nanometre-scale movements. Motor compliance may be an unavoidable consequence of this lever-arm design²³, but the compliance itself has important functional advantages. It allows force generation against an immobile cargo. It provides a sensor for modulation of the ATPase activity by the external load, as required for efficient energy transduction. Finally, as most motor proteins have more than one head domain attached to a single tail, compliance within the head allows two or more heads to work together. □

Methods

Fibre mechanics. Mechanical measurements on single fibres from anterior tibialis muscles of *Rana temporaria* were made at 3–4°C as described^{16,24}. Average sarcomere length in the ~2 mm segment in the X-ray beam was measured by a striation follower. Intact fibres were electrically stimulated for 1.5 s; after 0.3 s at constant length, 4,000 stretch-shortening cycles were applied (Fig. 1a); after the first 10 ms of imposed length oscillation the force response was the same in each cycle, whether the oscillation began with a stretch or release^{10,25}. For rigor experiments (Fig. 2), isolated fibres were permeabilized in relaxing solution (5.4 mM Na₂ATP, 7.7 mM MgCl₂, 25 mM EGTA, 19.1 mM creatine phosphate, 1 mg ml⁻¹ creatine kinase, 20 mM butanedione monoxime (BDM), 100 mM TES, 10 mM reduced glutathione) containing α -toxin at 18°C for 20 min^{16,26}. Rigor was induced at zero force by incubating in rigor solution (34 mM EDTA, 100 mM TES) plus 20 mM BDM, for 10 min at 0°C^{20,27}. Fibres were stretched by ~4 nm per half-sarcomere at constant velocity for 1.5 s, 4,000 stretch-shortening cycles were imposed, then fibre length was slowly returned to its initial value.

X-ray diffraction. X-ray diffraction measurements were made at stations 2.1 and 16.1 of the CLRC Daresbury Laboratory, UK, with a monochromator/mirror X-ray camera and two-dimensional gas-filled detector²⁸. Fibre-to-detector distance was 2.3 or 3.1 m. X-ray data were analysed using the XOTOKO and BSL packages from Daresbury. Reflection intensities were measured after fitting and subtracting a linear background; the intensity of the M3 reflection (I_{M3}), which has a spacing of about 14.5 nm in active contraction and rigor, was calculated by integrating between reciprocal spacings of 1/12.5 and 1/17.6 nm⁻¹ axially and 1/91 nm⁻¹ on either side of the meridian in active intact fibres (Fig. 1) and 1/13.1, 1/15.9 nm⁻¹ and 1/130 nm⁻¹, respectively, in rigor (Fig. 2). There was no significant change in the width of the M3 reflection during the length-change cycle either along or across the meridian, in either condition. For precise measurements of reflection spacings, a polynomial background was subtracted and the centre of each reflection determined either by fitting a gaussian (Peakfit, Jandel Scientific) or from the centre-of-mass calculated with the program HV written by A. Stewart. The spacing of the 5.9-nm actin-based layer line was measured by integrating from 1/91 to 1/6.5 nm⁻¹ from the meridian; its intensity changed by only 5% during the length-change cycle, increasing during the stretch.

Structural model. The effects of myosin head deformation on I_{M3} were calculated for bending of the light-chain domain (Fig. 3) as a uniform rod clamped at residue 770 under an axial strain (z) at residue 843. Axial displacement of each atom, assumed to be parallel to the 829–843 vector, was calculated as $z^2(3L - x)x^2/2L^3$, where x is distance of the atom from residue 770 and L is the value of x at residue 829, the sharp bend in the long heavy-chain helix. Residues between 829 and 843 were assigned $x = L$. ATP-driven tilting of the light-chain domain was at residue 707.

I_{M3} was initially calculated from the 1/14.5 nm⁻¹ Fourier component of the axial mass projection of the myosin head, including every atom of both the heavy and light chains, assuming that the population of myosin heads that respond to a length step and bear the active force all have the same conformation and are responsible for the observed changes in I_{M3} (ref. 29). The calculations were repeated with different values for myosin head compliance in order to find the compliance required to reproduce the observed changes in I_{M3} . The single-conformation model is unrealistic in that the catalytic domains of myosin heads with head/rod junctions on a 14.5 nm axial repeat could not bind to actin sites with a 5.5 nm axial repeat. We calculated the effects of the mismatch in filament periodicities for an array of 50 myosin heads with their head-rod junctions (residue 843) fixed on a 14.5 nm repeat. In addition to the elastic strain and inter-domain tilt parameters used in the single-conformation model, the extra strain required for the catalytic domain to bind in its normal conformation to the nearest actin site was simulated by applying an additional strain, chosen at random in the range ± 2.75 nm, to the light-chain domain of each individual head. The Fourier transform in the region of 1/14.5 nm⁻¹ was then calculated for the axial mass projection of the 50-head array. The static values of I_{M3} were reduced by about one third compared to those from the single-conformation model, but the fractional changes in I_{M3} produced by the length-change protocols in Fig. 3b–d were essentially the same as in the single-conformation model. The latter gives an accurate estimate of fractional changes in I_{M3} for the protocols used here because the effect of conformational dispersion is constant during the synchronous conformational change produced by rapid length changes, as shown analytically for the case of three populations of heads with different conformations²².

Received 9 February; accepted 18 September 1998.

- Huxley, H. E. The mechanism of muscular contraction. *Science* **164**, 1356–1366 (1969).
- Huxley, A. F. Muscular contraction. *J. Physiol. (Lond.)* **243**, 1–43 (1974).
- Raymont, I. *et al.* Structure of the actin–myosin complex and its implications for muscle contraction. *Science* **261**, 58–65 (1993).
- Raymont, I. *et al.* Three-dimensional structure of myosin subfragment-1: a molecular motor. *Science* **261**, 50–58 (1993).
- Irving, M. *et al.* Tilting of the light chain region of myosin during step length changes and active force generation in skeletal muscle. *Nature* **375**, 688–691 (1995).
- Cooke, R. Actomyosin interaction in striated muscle. *Physiol. Rev.* **77**, 671–697 (1997).
- Cecchi, G., Griffiths, P. J. & Taylor, S. Muscular contraction: kinetics of cross-bridge attachment studied by high-frequency stiffness measurements. *Science* **217**, 70–72 (1982).
- Huxley, A. F. & Simmons, R. M. Proposed mechanism of force generation in striated muscle. *Nature* **233**, 533–538 (1971).
- Huxley, A. F. Muscle structure and theories of contraction. *Prog. Biophys. Biophys. Chem.* **7**, 255–318 (1957).
- Lombardi, V., Piazzesi, G. & Linari, M. Rapid regeneration of the actin-myosin power stroke in contracting muscle. *Nature* **355**, 638–641 (1992).

11. Huxley, H. E., Faruqi, A. R., Kress, M., Borda, J. & Koch, M. H. J. Time resolved X-ray diffraction studies of the myosin layer line reflections during muscle contraction. *J. Mol. Biol.* **158**, 637–684 (1982).
12. Lombardi, V. *et al.* Elastic distortion of myosin heads and repriming of the working stroke in muscle. *Nature* **374**, 553–555 (1995).
13. Huxley, H. E., Stewart, A., Sosa, H. & Irving, T. X-ray diffraction measurements of the extensibility of actin and myosin filaments in contracting muscle. *Biophys. J.* **67**, 2411–2421 (1994).
14. Wakabayashi, K. *et al.* X-ray diffraction evidence for the extensibility of actin and myosin filaments during muscle contraction. *Biophys. J.* **67**, 2422–2435 (1994).
15. Ford, L. E., Huxley, A. F. & Simmons, R. M. The relation between stiffness and filament overlap in stimulated frog muscle fibres. *J. Physiol. (Lond.)* **311**, 219–249 (1981).
16. Linari, M. *et al.* The stiffness of skeletal muscle in isometric contraction and rigor: the fraction of myosin heads bound to actin. *Biophys. J.* **74**, 2459–2473 (1998).
17. Cooke, R. & Franks, K. All myosin heads form bonds with actin in rigor rabbit skeletal muscle. *Biochemistry* **19**, 2265–2269 (1980).
18. Lovell, S. J., Knight, P. J. & Harrington, W. F. Fraction of myosin heads bound to thin filaments in rigor fibrils from insect flight and vertebrate muscles. *Nature* **293**, 664–666 (1981).
19. Yagi, N. *et al.* Small-angle X-ray diffraction of muscle using undulator radiation from the Tristram main ring at KEK. *J. Synchrotron Rad.* **3**, 305–312 (1996).
20. Higuchi, H., Yanagida, T. & Goldman, Y. E. Compliance of thin filaments in skinned fibers of rabbit skeletal muscle. *Biophys. J.* **69**, 1000–1010 (1995).
21. Piazzesi, G. *et al.* Changes in the X-ray diffraction pattern from single intact muscle fibres produced by rapid shortening and stretch. *Biophys. J.* **68** (suppl.), 92–96 (1995).
22. Irving, M., Lombardi, V., Piazzesi, G. & Ferenczi, M. A. Myosin head movements are synchronous with the elementary force-generating process in muscle. *Nature* **357**, 156–158 (1992).
23. Howard, J. & Spudis, J. A. Is the lever arm of myosin a molecular elastic element? *Proc. Natl Acad. Sci.* **93**, 4462–4464 (1996).
24. Lombardi, V. & Piazzesi, G. The contractile response during steady lengthening of stimulated frog muscle fibres. *J. Physiol. (Lond.)* **431**, 141–171 (1990).
25. Piazzesi, G., Linari, M., Reconditi, M., Vanzi, F. & Lombardi, V. Cross-bridge detachment and attachment following a step stretch imposed on active single frog muscle fibres. *J. Physiol. (Lond.)* **498**, 3–15 (1997).
26. Nishiyee, E., Somlyo, A. V., Török, K. & Somlyo, A. P. The effect of MgADP on crossbridge kinetics: a laser flash photolysis study of guinea pig smooth muscle. *J. Physiol. (Lond.)* **460**, 247–271 (1993).
27. Bershtinsky, S. *et al.* Mechanical and structural properties underlying contraction of skeletal muscle fibers after partial EDC cross-linking. *Biophys. J.* **71**, 1462–1474 (1996).
28. Towns-Andrews, E. *et al.* Time-resolved X-ray diffraction station: X-ray optics, detectors, and data acquisition. *Rev. Sci. Instrum.* **60**, 2346–2349 (1989).
29. Irving, M. & Piazzesi, G. Motions of myosin heads that drive muscle contraction. *News Physiol. Sci.* **12**, 249–254 (1997).
30. Merritt, E. A. & Bacon, D. J. Raster3D: Photorealistic molecular graphics. *Methods Enzymol.* **277**, 505–524 (1997).

Acknowledgements. We thank the non-crystalline diffraction and detector development teams at the CLRC Daresbury Laboratory for hardware and software for the X-ray measurements, A. Aiazzi and M. Dolfi for skilled mechanical and electronic assistance, and the UK MRC, EC, Intas, CNR (Italy), and Telethon (Italy) for financial support.

Correspondence should be addressed to M.I.

New pathway to polyketides in plants

Stefan Eckermann*, Gudrun Schröder*, Jürgen Schmidt†, Dieter Strack†, Ru A. Edrada‡, Yrjö Helariutta§, Paula Elomaa§, Mika Kotilainen§, Ilkka Kilpeläinen§, Peter Proksch‡, Teemu H. Teeri§ & Joachim Schröder*

* Institut für Biologie II, Universität Freiburg, Schänzlestrasse 1, 79104 Freiburg, Germany

† Institut für Pflanzenbiochemie, Weinberg 3, 06120 Halle (Saale), Germany

‡ Lehrstuhl für Pharmazeutische Biologie, Mittlerer Dallenbergweg 64, 97082 Würzburg, Germany

§ Institute of Biotechnology, Viikki Biocenter, PO Box 56, FIN-00014 University of Helsinki, Finland

The repertoire of secondary metabolism (involving the production of compounds not essential for growth) in the plant kingdom is enormous, but the genetic and functional basis for this diversity is hard to analyse as many of the biosynthetic enzymes are unknown. We have now identified a key enzyme in the ornamental plant *Gerbera hybrida* (Asteraceae) that participates in the biosynthesis of compounds that contribute to insect and pathogen resistance. Plants transformed with an antisense construct of *gchs2*, a complementary DNA encoding a previously unknown

function^{1,2}, completely lack the pyrone derivatives gerberin and parasorboside. The recombinant plant protein catalyses the principal reaction in the biosynthesis of these derivatives: GCHS2 is a polyketide synthase that uses acetyl-CoA and two condensation reactions with malonyl-CoA to form the pyrone backbone of the natural products. The enzyme also accepts benzoyl-CoA to synthesize the backbone of substances that have become of interest as inhibitors of the HIV-1 protease^{3–5}. GCHS2 is related to chalcone synthase (CHS) and its properties define a new class of function in the protein superfamily. It appears that CHS-related enzymes are involved in the biosynthesis of a much larger range of plant products than was previously realized.

CHS, stilbene synthase and acridone synthase are polyketide synthases that are found in plants; they share >65% identity. They use aromatic CoA-esters as a starter substrate, perform three condensation reactions with malonyl-CoA, and form new aromatic ring systems⁶. We have previously characterized cDNAs encoding two CHS enzymes from *G. hybrida* and a cDNA encoding a CHS-like protein (GCHS2) with 73% identity to the CHS enzymes. In contrast to CHS, GCHS2 is highly expressed in almost all tissues of the plant. Its presence does not correlate with flavonoid biosynthesis¹, and its substrate preferences differ from those of the CHSs^{1,2}.

By using antisense transformation of *G. hybrida* with the *gchs2* cDNA, we obtained 14 plant lines that lacked *gchs2* expression partially or completely⁷. We next analysed methanol extracts from leaf, floral stem and corolla tissue of four transgenic lines, and from non-transgenic plants. This analysis showed that downregulation of *gchs2* correlates with the absence of a dominant substance found in control plants (Fig. 1a, b). In non-transgenic plants, the ontogenetic distribution of this substance correlates with the presence of *gchs2* messenger RNA (found in large amounts in all tissues except anthers), but not with the transcripts of the other CHS-related genes. A partial suppression of CHS genes takes place in the

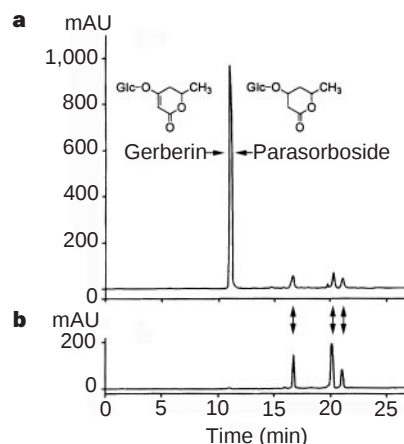


Figure 1 HPLC analysis of methanol extracts from leaves. **a**, Wild-type plant. **b**, *gchs2* antisense transformant, with no detectable *gchs2* expression. The major peak in **a** contains a mixture of gerberin and parasorboside. The peaks marked with arrows in **b** are hydroxycinnamic acid derivatives which are present at higher concentrations in the transgenic plants. A second independent transformant with no detectable *gchs2* mRNA also showed the pattern seen in **b**. Two more transformants with partial antisense effect (17% and 97% of wild-type *gchs2* mRNA was detected) showed a reduced gerberin/parasorboside peak (17% and 72% of the wild-type, respectively). Corolla samples from antisense plants showed the same pattern of change as leaves. The wavelength used for detection of the pyrone derivatives (254 nm) does not detect most flavonoids, but analyses at the appropriate wavelength indicate that their concentration is affected only slightly (<10% reduction) in *gchs2* antisense transformants. HPLC was done with an RP18 column, a linear gradient from 100% phosphoric acid (pH 2) to 100% methanol in 30 min, and a flow rate of 1 ml min⁻¹. mAU, absorption units ×10⁻³.

|| Present address: The Lewis B. and Dorothy Cullman Program for Molecular Systematics Studies, The New York Botanical Garden, Bronx, New York 10458-5126, USA.

strongest anti-*gchs2* lines, as shown by the slight reduction in colour of these plants⁷, but the flavonoid level remains ~90% that of non-transgenic plants.

NMR and mass spectroscopic analysis showed that the dominant peak from the control plants is a mixture of two pyrone glucosides (Fig. 1a), namely gerberin, which has been isolated previously from *Gerbera*⁸, and parasorboside, which was not previously known from this plant, but has been described from cranberry⁹ and the fern *Osmunda japonica*¹⁰ and is present in large amounts in berries of mountain ash (*Sorbus aucuparia*)^{11,12}. The biosynthesis of these substances has not been studied before.

The biological activities of parasorboside and its immediate derivatives are well known. The glucoside is a major cause of the bitter taste of mountain ash berries¹² and acts to prevent yellow butterfly larvae from feeding¹⁰. Parasorboside is the precursor of parasorbic acid¹¹, which inhibits the growth of fungi¹³ and bacteria¹⁴; it also delays the germination of seeds¹⁵ and potato tubers¹⁶. Studies of the pathogen susceptibility of anti-*gchs2* lines indicate that gerberin, parasorboside or their derivatives may also protect *Gerbera* from fungal pathogens and attack by insects (P.E., S. Remes and T.H.T., unpublished data).

The correlation between *gchs2* expression and the presence of the two pyrone glucosides indicates that GCHS2 may be involved in their biosynthesis, and the similarity of GCHS2 to CHS enzymes is consistent with the hypothesis that GCHS2 is a polyketide synthase. The aglycone backbone of both gerberin and parasorboside could be synthesized by a polyketide synthase reaction, but we believe that the enzyme has properties very different from those of CHS enzymes (Fig. 2a). We predict that, in contrast with CHSs, GCHS2 uses acetyl-CoA as a starter substrate instead of phenylpropanoid CoA-esters, performs two condensation reactions with malonyl-CoA (instead of three) and releases 6-methyl-4-hydroxy-2-pyrone

(methylpyrone) (Fig. 2b). This model includes cyclization to the pyrone, but does not exclude non-enzymatic ring closure, which has been observed in other polyketide synthases if the reactions are terminated after the second condensation^{17–20}.

We investigated the function of GCHS2 after recloning the *gchs2* cDNA for protein expression in *Escherichia coli*, and purifying the recombinant protein. The thin-layer chromatography (TLC) radio-scan in Fig. 3a shows the product obtained from acetyl-CoA and radioactive malonyl-CoA. This product was identified by liquid chromatography (LC)/mass spectrometry (MS) and gas chromatography (GC)/electron impact mass spectrometry (EIMS) as methylpyrone (see Methods), as predicted (Fig. 2). Unexpectedly, control incubations containing only [¹⁴C]malonyl-CoA gave the same product (Fig. 3b), but at a rate of incorporation of ¹⁴C that was two- to threefold lower than in the presence of the predicted starter CoA-ester. It was, therefore, possible that acetyl-CoA unspecifically increased the incorporation from malonyl-CoA. Experiments with [¹⁴C]acetyl-CoA and unlabelled malonyl-CoA, however, showed that the radioactive product was formed, indicating that acetyl-CoA is used as a substrate. Acetoacetyl-CoA, the predicted intermediate, was a ~2.5-fold more efficient substrate than acetyl-CoA. The product was methylpyrone, as shown by LC/MS and GC/EIMS.

The enzyme could form methylpyrone by decarboxylation of [¹⁴C]malonyl-CoA to produce acetyl-CoA, which is then used as the starter substrate. Measurements of [¹⁴C]acetyl-CoA confirmed that GCHS2 efficiently decarboxylates malonyl-CoA, with rates that are tenfold higher than the rate of formation of methylpyrone. Adding acetyl-CoA reduced the decarboxylase activity, but it still remained fivefold higher than the condensing reactions. Self-priming by chain-extender decarboxylation is known from other polyketide synthases with no structural similarities to the CHS-related proteins, such as the fungal 6-methylsalicylic acid synthase²¹.

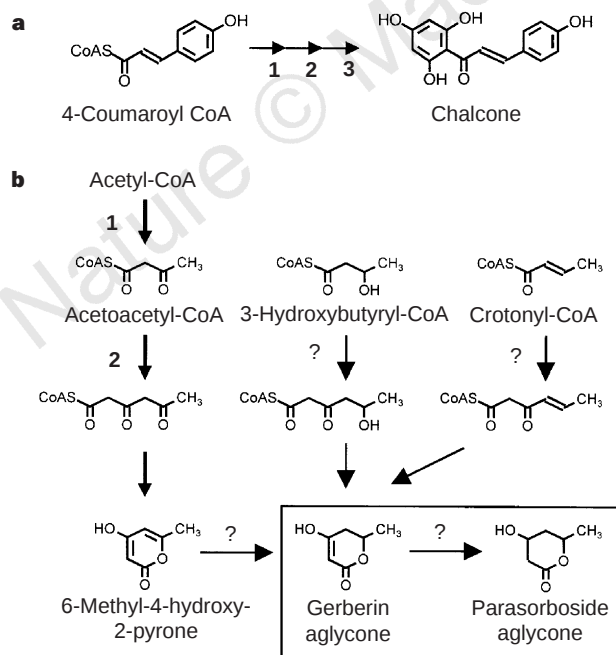


Figure 2 Biosynthetic pathways. **a**, The CHS-catalysed reaction. **b**, Proposed biosynthesis of the aglycone backbone of gerberin and parasorboside by a polyketide synthase reaction (bold arrows, at left). The question marks indicate where the reduction steps in the biosynthesis of the natural products could occur. Numbers in **a**, **b** indicate the condensation reactions.

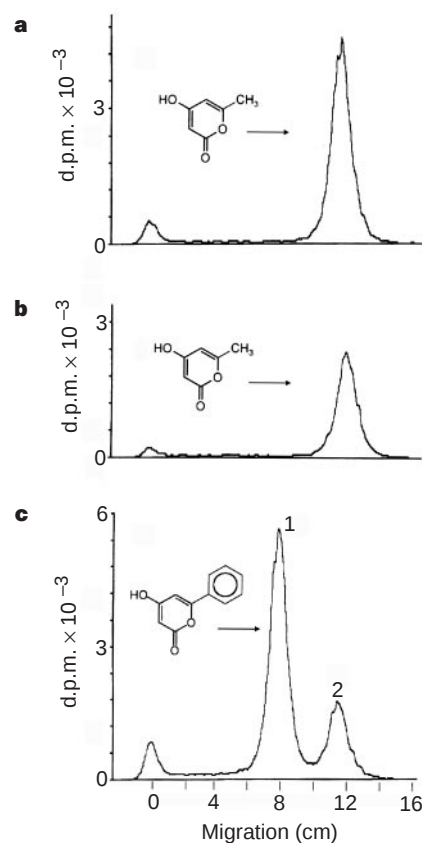


Figure 3 Thin-layer-chromatographic analysis of radioactive products synthesized by GCHS2 *in vitro*. Substrates are: **a**, acetyl-CoA plus [¹⁴C]malonyl-CoA; **b**, [¹⁴C]malonyl-CoA; **c**, benzoyl-CoA plus [¹⁴C]malonyl-CoA.

The formation of the gerberin and parasorboside aglycones requires one and two reductions in the pyrone ring, respectively. It was possible that the first reducing step occurs at the level of the C4 intermediates and that 3-hydroxybutyryl-CoA or crotonyl-CoA could lead directly to the gerberin aglycone (Fig. 2b). However, assays with these substrates and [14 C]malonyl-CoA showed the same low rate of incorporation as is found with malonyl-CoA only, and high-performance liquid chromatography (HPLC) confirmed that the *in vitro* product is methylpyrone, supporting the idea that the reduction step does not occur at the level of C4 intermediates.

The activity of GCHS2 with benzoyl-CoA^{1,2} led us to analyse the products synthesized in the presence of benzoyl-CoA and [14 C]malonyl-CoA (Fig. 3c). Product 2 is the same as that detected from incubations with only malonyl-CoA (Fig. 3b), indicating that it is formed by the same mechanism. The dominant product 1 was identified by LC/MS and GC/EIMS as 6-phenyl-4-hydroxy-2-pyrone (phenylpyrone; see Methods), a compound predicted to be produced by a pyrone synthase that is active with benzoyl-CoA. The material at the origin of the chromatogram in Fig. 3c probably represents polymerization products of the unstable pyrones. Phenylpyrone derivatives are not known from *Gerbera*, but they form the aglycone backbone in the psilotsins of the Psilotaceae²². The enzyme also accepts other small hydrophobic starter CoA-esters: the activities with propionyl-CoA and isovaleryl-CoA are comparable to those with acetyl-CoA, and the activities with butyryl-CoA are approximately twofold higher than those with acetyl-CoA. The products are all 4-hydroxy-2-pyrones, as predicted, but neither they nor their derivatives are known from *Gerbera*. 4-Coumaroyl-CoA, the preferred substrate of CHSs, is not accepted by the enzyme.

Figure 2 shows the most important differences between the CHS enzymes involved in flavonoid biosynthesis and GCHS2, which is unique in its specialization for only two condensation reactions, its complete lack of activity with phenylpropanoid CoA-esters, and its high activity with acetyl-CoA, the only substrate that leads to formation of the metabolites accumulating in *Gerbera*. CHSs also show a broad substrate acceptance *in vitro*. The CHS from parsley²³, for example, uses butyryl-CoA, hexanoyl-CoA, and benzoyl-CoA with almost the same efficiency as 4-coumaroyl-CoA. The CHSs from *Pinus strobus* and *Sinapis alba* show similar properties to parsley CHS (J. C. et al., unpublished data). A recent study with purified *Pinus sylvestris* (Scots pine) CHS and isovaleryl-CoA or isobutyryl-CoA showed that the enzyme synthesized pyrones as well as the products expected from three condensation reactions²⁴. Acetyl-CoA, however, is a poor substrate for CHSs (compared with the amount of 4-coumaroyl-CoA used, only <10% acetyl-CoA is used)²³. In contrast to GCHS2, incubations of CHSs with only [14 C]malonyl-CoA do not lead to significant product formation, although CHSs can decarboxylate malonyl-CoA²⁵.

Our results show that the CHS-like protein in *Gerbera* is a pyrone synthase; accordingly we rename GCHS2 as 2-pyrone synthase (2PS) and rename the gene *gchs2* as *g2ps1* (*Gerbera* contains a small family of related pyrone synthase genes²). The properties of 2-pyrone synthase indicate that it might be an evolutionarily old form, as its function is more similar to that of fungal and bacterial polyketide synthases than to that of the other members of the family of CHS-related enzymes. However, the similarity of 2-pyrone synthase to the CHSs in *Gerbera* and the sporadic distribution of gerberin-type compounds lead us to believe that the enzyme evolved by CHS gene duplication and mutation. Stilbene synthases are proposed to have evolved from CHSs by the same mechanism²⁶. The family of CHS-related proteins contains enzymes with surprisingly diverse functions, and it seems that relatively small differences in the proteins are responsible for the substrate specificities and the progressivity of the condensation reactions. A better understanding of these mechanisms could allow the design of enzymes for the production of new, useful substances in metabolically altered hosts. □

Methods

Materials. [14 C]malonyl-CoA and [14 C]acetyl-CoA (both 55 mCi mmol⁻¹) were obtained from Biotrend. The other CoA-esters were from the laboratory collection or purchases from Aldrich Biochemicals. The gerberin aglycone (5,6-dihydro-6-methyl-4-hydroxy-2-pyrone) was from Aldrich Biochemicals.

Protein expression in *E. coli*. *gchs2* (*g2ps1*) was recloned from the original cDNA clone pGC11.4 (ref. 1) into the *Sall* and *Bam*HI sites of pQE-6(*Not*I)²⁶. We used a *Xho*I site a few base pairs upstream of the 5' start codon (AUG) and a *Bam*HI site from the vector at the 3' end. The cloning procedure added six amino acids (MGASSK) at the protein's amino terminus; these amino acids have no effect on substrate specificity¹. The enzyme was expressed in *E. coli* strain RM82 containing plasmid pUBS520 (ref. 27).

Enzyme purification. The collected cells were resuspended in 0.05 M HEPES buffer (pH 7.5), broken in a French pressure cell (1,018 p.s.i.), and centrifuged to remove insoluble material (20,000g at 4 °C for 20 min). The purification used the following steps, with all buffers containing 2 mM dithiothreitol: first, fractionated ammonium sulphate precipitation, with most of the enzyme in the 40–60% fraction; second, size-exclusion chromatography (Fractogel EMD BioSEC S, Merck); and third, anion-exchange chromatography (Fractogel EMD DEAE 650 S, Merck) with a step gradient from 0 to 1 M NaCl. The enzyme eluted at 90 mM NaCl. The final preparation was at least 90% pure, as analysed by SDS gel electrophoresis.

Pyrone synthase assay. The incubations (0.1 ml) contained 10 μ M or 15 μ M starter CoA-ester, 15 μ M [14 C]malonyl-CoA (73,000 d.p.m.), 0.1 M HEPES (pH 6), and 2–5 μ g enzyme. They were stopped after 30 min at 30 °C by extraction of the products into ethyl acetate. The radioactive products were quantified by TLC²⁸. The amount of enzyme and the incubation time were chosen to produce the incorporation of $\leq 30\%$ of the radioactive substrate.

Malonyl-CoA decarboxylase assay. We used the same conditions as for the pyrone synthase assay. The aqueous phase of the ethyl acetate extraction was dried under vacuum. The residue was dissolved in water and analysed on silica-60 TLC plates. The separation of malonyl-CoA (R_f 0.27) and acetyl-CoA (R_f 0.7) was in isopropanol:H₂O:25% aqueous ammonia (80:5:13 v/v/v). The quantification was as described for the products of the condensing reactions.

NMR spectroscopy and mass spectroscopy. The structures of gerberin and parasorboside were determined using ¹H–¹H correlated (2D COSY and TOCSY) and ¹H–¹³C correlated (2D HMQC and HMBC) NMR techniques. NMR spectra are available from the authors on request. Mass spectra (ESI-MS) were recorded on a Finnigan MAT TSQ-7000 triple-stage quadrupole mass spectrometer. The temperature of the heated capillary (20 V) was 200 °C and the electrospray capillary voltage was set to 3.5 kV. Nitrogen was used both as the sheath (70 p.s.i.) and as the auxiliary gas and argon as the collision gas.

LC/negative ion electrospray mass spectra and data. We used a Finnigan MAT TSQ 7000 instrument (electrospray voltage 3.5 kV, APICID offset voltage 10 eV, heated capillary temperature 220 °C, sheath gas nitrogen) coupled with a Micro-Tech Ultra-Plus MicroLC system equipped with a RP18-column (4 μ m, 1 $\times$ 100 mm, SEPSERV). HPLC conditions were as follows: gradient system starting from 20% CH₃CN in H₂O (each contained 0.2% acetic acid) to 90% CH₃CN in H₂O within 10 min; flow rate 70 μ l min⁻¹. All relative retention times (RR_i) were calculated with respect to the retention time of 3-methyl-6-styryl-4-hydroxy-2-pyrone (R_t = 10.97 min). All mass spectra were averaged and background-subtracted. Data obtained are relative retention times and, for the negative ions, m/z (mass/charge ratio) and relative intensity values (relative intensity is shown in parentheses): 6-phenyl-4-hydroxy-2-pyrone: RR_i = 0.8, [M-H]⁻, 187 (100), [M-H-CO₂]⁻ 143 (60); 6-methyl-4-hydroxy-2-pyrone (authentic sample available): RR_i = 0.87, [M-H]⁻ 125 (100), [M-H-CO₂]⁻ 81 (10).

GC/EIMS of trimethylsilyl ethers (TMSi) and data. The GC/MS measurements were carried out with an MD 800 (Fisons Instruments) under the following conditions: electron ionization (70 eV), capillary column DB-5MS (15 m $\times$ 0.32 mm, 0.25 μ m film thickness), source temperature 200 °C, injection temperature 250 °C, interface temperature 300 °C, carrier gas He, flow rate 1.3 ml min⁻¹, splitless injection, temperature program: 3 min at 80 °C, increased to 290 °C at 10 °C min⁻¹, hold at 290 °C for 10 min. The relative retention times were calculated with respect to squalene (R_t = 18.94 min). Data obtained from the ion chromatograms are: 6-phenyl-4-hydroxy-2-pyrone: RR_i 0.70, 260 (M⁺, 54), 245 ([M-Me]⁺, 23), 232 ([M-CO]⁺, 100), 203 (8), 158 (14), 105 (C₆H₅-

CO⁺, 53), 77 (C₆H₅⁺, 60), 73 (TMSi⁺, 84); 6-methyl-4-hydroxy-2-pyrone: RR_t 0.35, 198 (M⁺, 18), 183 ([M-Me]⁺, 16), 170 ([M-CO]⁺, 54), 155 ([M-CO-Me]⁺, 15), 139 ([M-Me-CO₂]⁺, 10), 127 ([M-Me-2CO]⁺, 13), 99 (12), 84 (13), 73 (TMSi⁺, 100), 43 (CH₃CO⁺, 55). The numbers show *m/z* values, and the key fragments and their relative intensities are indicated in parentheses.

Received 4 August; accepted 14 October 1998.

- Helariutta, Y. *et al.* Chalcone synthase-like genes active during corolla development are differentially expressed and encode enzymes with different catalytic properties in *Gerbera hybrida* (Asteraceae). *Plant Mol. Biol.* **28**, 47–60 (1995).
- Helariutta, Y. *et al.* Duplication and functional divergence in the chalcone synthase gene family of Asteraceae: evolution with substrate change and catalytic simplification. *Proc. Natl Acad. Sci. USA* **93**, 9033–9038 (1996).
- Thaisrivongs, S. *et al.* Structure-based design of HIV protease inhibitors: 5,6-dihydro-4-hydroxy-2-pyrone as effective, nonpeptidic inhibitors. *J. Med. Chem.* **39**, 4630–4642 (1996).
- Hagen, S. E. *et al.* Synthesis of 5,6-dihydro-4-hydroxy-2-pyrone as HIV-1 protease inhibitors: the profound effect of polarity on antiviral activity. *J. Med. Chem.* **40**, 3707–3711 (1997).
- Tait, B. D. *et al.* 4-hydroxy-5,6-dihydropyrone. 2. Potent non-peptide inhibitors of HIV protease. *J. Med. Chem.* **40**, 3781–3792 (1997).
- Schröder, J. A family of plant-specific polyketide synthases: facts and predictions. *Trends Plant Sci.* **2**, 373–378 (1997).
- Elomaa, P., Helariutta, Y., Kotilainen, M. & Teeri, T. H. Transformation of antisense constructs of the chalcone synthase gene superfamily into *Gerbera hybrida*: differential effect on the expression of family members. *Mol. Breed.* **2**, 41–50 (1996).
- Nagumo, S., Toyonaga, T., Inoue, T. & Nagai, M. New glucosides of a 4-hydroxy-5-methylcoumarin and a dihydro- α -pyrone from *Gerbera Jamesonii* hybrida. *Chem. Pharm. Bull. (Tokyo)* **37**, 2621–2623 (1989).
- Cardellina, J. G. II & Meinwald, J. Isolation of parasorbic acid from the cranberry plant, *Vaccinium macrocarpon*. *Phytochemistry* **19**, 2199–2200 (1980).
- Numata, A. *et al.* Plant constituents biologically active to insects. VI. Antifeedants for larvae of the yellow butterfly *Eurema hecabe mandarina* in *Osmunda japonica*. *Chem. Pharm. Bull. (Tokyo)* **38**, 2862–2865 (1990).
- Tschesche, R., Hoppe, H.-J., Snatzke, G., Wulff, G. & Fehlbauer, H.-W. Über Parasorbosid, den glykosidischen Vorläufer der Parasorbinsäure, aus Vogelbeeren. *Chem. Ber.* **104**, 1420–1428 (1971).
- Pyysalo, H. & Kuusi, T. Phenolic compounds from the berries of mountain ash, *Sorbus aucuparia*. *J. Food Sci.* **34**, 636–638 (1974).
- Buston, H. W. & Roy, S. K. The physiological activity of some simple unsaturated lactones. I. Effect on the growth of certain microorganisms. *Arch. Biochem.* **22**, 1–7 (1949).
- Kuhn, R., Jerchel, D., Moewus, F., Möller, E. F. & Lettré, H. Über die chemische Natur der Blastokoline und ihre Einwirkung auf keimende Samen, Pollenkörner, Hefen, Bakterien, Epithelgewebe und Fibroblasten. *Naturwissenschaften* **37**, 463 (1943).
- Oster, U., Bloss, I. & Rüdiger, W. Natural inhibitors of germination and growth. IV. Compounds from fruit and seeds of mountain ash (*Sorbus aucuparia*). *Z. Naturforsch.* **42c**, 1179–1184 (1987).
- Moewus, F. & Schader, E. Die Wirkung von Cumarin und Parasorbinsäure auf das Austreiben von Kartoffelknollen. *Z. Naturforsch.* **6b**, 112–115 (1951).
- Yalpani, M., Willecke, K. & Lynen, F. Triacetic acid lactone, a derailment product of fatty acid biosynthesis. *Eur. J. Biochem.* **8**, 495–502 (1969).
- Dimroth, P., Walter, H. & Lynen, F. Biosynthese von 6-Methylsalicylsäure. *Eur. J. Biochem.* **13**, 98–110 (1970).
- Kurosaki, F., Kizawa, Y. & Nishi, A. Derailment product in NADPH-dependent synthesis of a dihydroisocoumarin 6-hydroxymellein by elicitor-treated carrot cell extracts. *Eur. J. Biochem.* **185**, 85–89 (1989).
- Pieper, R., Luo, G., Cane, D. E. & Khosla, C. Cell-free synthesis of polyketides by recombinant erythromycin polyketide synthases. *Nature* **378**, 263–266 (1995).
- Dimroth, P., Ringelmann, E. & Lynen, F. 6-Methylsalicylic acid synthetase from *Penicillium patulum*. Some catalytic properties of the enzyme and its relation to fatty acid synthetase. *Eur. J. Biochem.* **68**, 591–596 (1976).
- McInnes, A. G., Yoshida, S. & Towers, G. H. N. A phenolic glycoside from *Psilotum nudum* (L.) Griseb. *Tetrahedron* **21**, 2939–2946 (1965).
- Schüz, R., Heller, W. & Hahlbrock, K. Substrate specificity of chalcone synthase from *Petroselinum hortense*. Formation of phloroglucinol derivatives from aliphatic substrates. *J. Biol. Chem.* **258**, 6730–6734 (1983).
- Zuurbier, K. W. M. *et al.* 4-Hydroxy-2-pyrone formation by chalcone and stilbene synthase with nonphysiological substrates. *Phytochemistry* (in the press).
- Kreuzaler, F., Light, R. J. & Hahlbrock, K. Flavanone synthase catalyzes CO₂ exchange and decarboxylation of malonyl-CoA. *FEBS Lett.* **94**, 175–178 (1978).
- Tropf, S., Lanz, T., Rensing, S. A., Schröder, J. & Schröder, G. Evidence that stilbene synthases have developed from chalcone synthases several times in the course of evolution. *J. Mol. Evol.* **38**, 610–618 (1994).
- Schenk, P. M., Baumann, S., Mattes, R. & Steinbiss, H.-H. Improved high-level expression system for eucaryotic genes in *Escherichia coli* using T7 RNA polymerase and rare tRNAs. *Biotechniques* **19**, 196–200 (1995).
- Lanz, T., Tropf, S., Marner, F.-J., Schröder, J. & Schröder, G. The role of cysteines in polyketide synthases: site-directed mutagenesis of resveratrol and chalcone synthases, two key enzymes in different plant-specific pathways. *J. Biol. Chem.* **266**, 9971–9976 (1991).

Acknowledgements. We thank T. J. Simpson for a sample of authentic 6-methyl-4-hydroxy-2-pyrone, L. Britsch for Fractogel EMD Butyl 650 S, and R. Mattes for *E. coli* strain RM82 with plasmid pUBS520. The group in Freiburg was supported by a grant from the Deutsche Forschungsgemeinschaft. T.H.T. and

P.P. thank the Academy of Finland and the Deutsche Forschungsgemeinschaft, respectively, for financial support.

Correspondence and requests for materials should be addressed to J.S. (e-mail: jschroe@ruf.uni-freiburg.de).

erratum

Crystal structure of the complex of the cyclin D-dependent kinase Cdk6 bound to the cell-cycle inhibitor p19^{INK4d}

Deborah H. Brotherton, Venugopal Dhanaraj, Scott Wick, Leonardo Brizuela, Peter J. Dommelle, Elena Volyanik, Xu Xu, Emilio Parisini, Brian O. Smith, Sharon J. Archer, Manuel Serrano, Stephen L. Brenner, Tom L. Blundell & Ernest D. Laue

Nature **395**, 244–250 (1998)

In the text we refer to the work of Batchelor *et al.* (ref. 25) on the structure of a GABP α / β -DNA complex, but GABP was incorrectly written as GABA (γ -aminobutyric acid).

In addition, some spurious dark patches were introduced into Fig. 1 during the production process but these do not affect the information conveyed by the figure. □

retraction

Identification and role of adenylyl cyclase in auxin signalling in higher plants

Takanari Ichikawa, Yoshihito Suzuki, Inge Czaja, Carla Schommer, Angela Leßnick, Jeff Schell & Richard Walden

Nature **390**, 698–701 (1997)

Some of the results reported in this Letter cannot be reproduced. We know that the data on protoplast division described for Figs 1a, c, 2a, b and 4, and the corresponding experimental procedures described in the Methods section and the text, are wrong. In fact, the data showing that cAMP can stimulate protoplast division in the absence of auxins are not correct. Hence, we need to retract this paper. We apologize for any misunderstanding that this might have caused.

Note from the Editor: One author, R.W., although concerned with the accuracy of parts of this paper, reserves judgement concerning its retraction and awaits the outcome of further experimentation. □

Molecular synthesis and amplification

New products covered here include a cDNA selection kit, dye—polymerase and high-efficiency polymerase formulations, separation systems for DNA synthesis, single-tube PCR and RT-PCR amplification systems.

ClonCapture

From Clontech

A new cDNA selection kit for the rapid capture of full-length cDNA clones

ClonCapture is a gene isolation system designed to enrich target clones from any plasmid library without extensive screening. Together with a partial cDNA probe, full-length target cDNA clones can be isolated in just two or three days, saving significant time over other methods, says Clontech. The strategy uses a ClonCapture RecA protein, a 200–600-bp biotinylated probe that binds homologous cDNA clones, which are then captured by streptavidin-coated magnetic beads. This kit is claimed to be more efficient than non-RecA methods, more reliable than techniques that require conversion to single-stranded libraries and more specific than methods using shorter probes.

Reader Enquiry No. 100

Pre-packed CPG columns

From Alpha DNA

DNA synthesis performed on a controlled-pore, glass (CPG) column

A system wherein the last nucleoside covalently attaches by its 3'-OH group to a solid support. The other bases of the oligonucleotide are added consecutively, and DNA synthesis advances from the 3' to the 5'-end (contrary to DNA synthesis in the living cell). CPG controlled-pore glass is a non-swelling rigid, porous matrix. The particles are about 150 µm in size and come in different porosities: 500 Å for oligonucleotides less than 55 bases long, 1,000 Å for up to 100–150 bases and 2,000 Å for more than 100 bases. The company's prepacked columns have of porosity 1,000 Å and could be used on the majority of commercially available DNA synthesizers, for example, Expedite, ABI, Biosearch, Millipore, Pharmacia (for Beckman instruments an adapter is required). The columns come in three sizes: 50 nmol, 200 nmol and 1.0 µmol. Other sizes are available upon request.

Reader Enquiry No. 101

µ.Macs

From Miltenyi Biotec

A new magnetic separator designed for molecular biology applications

Designed for the magnetic isolation of molecules, for example mRNA, the separator can hold up to four micro-columns, allowing up to four isolations to be performed in parallel.



Micro-column separation with µ.Macs.

In combination with µMacs mRNA isolation kits, these columns can be used to isolate highly pure mRNA from cells, tissue or total RNA. Therefore, the mRNA is magnetically labelled using Macs Oligo (dT) microbeads and subsequently captured on a micro-column that is placed in the magnetic field. Genomic DNA, proteins and ribosomal RNA are effectively washed away and the highly pure mRNA is subsequently eluted from the column. With each column, up to 10 µg mRNA can be isolated in less than 15 minutes. The technology should help provide pure, full-length mRNA with high yields.

Reader Enquiry No. 102

Amplifications

TaKaRa Z-Taq

From Takara Shuzo

TaKaRa Z-Taq DNA polymerase was developed to achieve rapid PCR

This polymerase is intended to combine PCR productivity that is fivefold faster than previous polymerases, and with processivity and sensitivity that is as high as TaKaRa's Ex Taq. It is said to provide high-throughput, multi-sample processing on a single cyclor. When handling 10,000 samples of 1,000 bp, for example, Z-Taq is stated to complete amplification of 1,000 bp in approximately 20 min — all PCR reactions will finish in about 1.5 days. **Reader Enquiry No. 103**

REDTaq

From Sigma

A new Taq formulation with a visible red layer at the bottom and standard enzyme performance

The colour eliminates the uncertainty that can occur with interruptions in pipetting and makes it easy to confirm proper mixing of samples. When the red colour is uniform a solution is thoroughly mixed. Moreover, REDTaq requires no loading buffers as the post-PCR samples are dense enough to be loaded directly onto an agarose gel. The red dye works as a tracking dye, migrating just faster than bromophenol blue. Intended to make accurate pipetting easier, this formu-

lation is provided in a concentration of one unit per microlitre. A uniform amount of enzyme is added to ensure consistency in the amount of end product.

Reader Enquiry No. 104

Hy-Tek ExpeRT RT-PCR

From Hybaid

Efficient RT-PCR in a one-tube format

This kit uses a new buffer system in which RT-PCR fragments of up to 6 kb can be achieved in a single step. The buffer composition is also said to reduce problems with secondary structures in RNA templates and to increase the specificity of primers. In addition, the kit will perform both conventional and *in situ* RT-PCR. The system is based on AMV reverse transcriptase and the company's ProofSprinter (*Taq/Pwo*) enzyme mix, allowing high fidelity RT-PCR with more than a stated threefold error reduction when compared to conventional methods.

Reader Enquiry No. 105

LightCycler

From Roche Diagnostics

Designed to complete a typical PCR experiment in <30 min and to offer online quantification of amplification products

This system combines rapid thermal cycling with an integrated microvolume fluorimeter, which performs measurements for each cycle in just 20 ms. The combination helps to eliminate the possibility of contamination, as amplification and detection occur in the same closed tube. The system offers users new analytical opportunities, such as the determination of melting points and the detection of mutations. The online detection system eliminates the need



Move over Tron: Roche's has a new Light Cycle.

for multiple samples and repeated removal of aliquots for analysis, where contaminants could be introduced. It also replaces the multiple handling steps of conventional post-PCR processing. The LightCycler system is based on fluorescence, resonance-energy transfer (FRET) detection. When two differently labelled 3' and 5' oligonucleotides are included with the conventional PCR reagents, the LightCycler excites the fluoroscein label on the 3' probe in the amplification product, which, in turn, excites the LC Red640 label attached to the 5' probe. Detection of this emission allows the quantification of the amplification product and will detect mutations through changes in the melting characteristics of the hybridization probes. The system will also detect DNA binding dyes, such as SYBR Green I, which secure newly synthesized dsDNA during PCR and emit a signal proportional to the DNA concentration.

Reader Enquiry No. 106

Synthesis systems

ABI 3948

From Perkin-Elmer

Nucleic acid synthesis and purification system

This system is designed to provide fully automated synthesis, cleavage, deprotection, purification and quantification of 48 primer-length oligonucleotides in about 24 h. It achieves this with approximately one half hour of hands-on time for instrument set up, automating just about all tasks from synthesis to sample collection. There is said to be no need for user intervention, and the oligos are ready to use in experiment-compatible 20 per cent acetonitrile.

Reader Enquiry No. 107

OligoPilot II DNA/RNA synthesizer

From Amersham Pharmacia Biotech

A pump-driven, benchtop oligonucleotide synthesizer

This system is designed for production at intermediate and pilot scales, and methods development and optimization at large scale. It is controlled by Unicorn software and equipped with five fixed-volume, flow-through steel column reactors with a volume range of 1.2–48 ml, and one adjustable-volume column reactor (30–150 ml). Scale up is said to be simple and reproducible from 0.01 to 5 mmol.

Reader Enquiry No. 108

UltraMild DNA synthesis reagents

From Glen Research

Biotin and fluoroscein are now available as β -cyanoethyl (CE) phosphoramidites

Labels that are currently available as CE phosphoramidites have one common property: they must be stable to the strongly

alkaline conditions required for the removal of the base protecting groups. This property is lacking in several interesting dyes and labels. Glen Research has been seeking an alternative protecting scheme for the normal CE phosphoramidites, which would allow UltraMild deprotection and would not react with a wider variety of tags and labels. A set of monomers using Pac-protected (phenoxyacetyl) dA and iPr-Pac-protected dG (4-isopropyl-phenoxyacetyl), along with acetyl protected dC, have met the desired criteria for deprotection. According to the manufacturer, cleavage and deprotection

can now be carried out in 2 h at room temperature with ammonium hydroxide or 0.05 M potassium carbonate in methanol. Isobutryl (ibu) protected dA can be deprotected under very mild conditions but its primary role is in photochemical procedures, where aromatic protecting groups can not be used. The ibu-protected dA can be used with the Ac-dC and ibu-dG.

Reader Enquiry No. 109

These notes are compiled by Brendan Horton from information provided by the manufacturers. For more details, fill in the reader service card bound inside the journal.

ADVERTISEMENTS

BECKMAN COULTER

FULLY AUTOMATED DNA SEQUENCER

www.beckmancoulter.com/ceq20

READER ENQUIRY NO. 47

Check out our homepage at <http://www.resgen.com>

BACs
YACs
PACs
cDNAs

Research Genetics provides clones and colony membranes for a number of human, mouse, and rat libraries as well as a custom screening service for the libraries.

Research Genetics, Inc.
U.S. or Canada 800-533-4363
U.K. 0-800-89-1393
FAX 256-536-9016

READER ENQUIRY NO. 30

Detect residual DNA with Wako's DNA Extractor Kit or Isolate Genomic DNA with DNA Extractor WB Kit

Both use the Nal Method
No Phenol No Chloroform

Wako

Wako BioProducts
Wako Chemicals USA, Inc.
1600 Bellwood Road
Richmond, VA 23237
Telephone: (804) 271-7677
Facsimile: (804) 271-7791
(800) 992-9256
bioproducts@wakousa.com

READER ENQUIRY NO. 39

Homogenizers?

- Tissue-Tearor™
- BeadBeater™
- Mini-BeadBeater™
- BioHomogenizer™

Check Us Out

www.biospec.com

BioSpec Products, Inc.
Ph/Fax 918-336-3363
e-mail info@biospec.com

New Probe Sizes
0.45cm

READER ENQUIRY NO. 54

Severn Biotech Ltd

Customer Services

Comprehensive Custom Service for DNA, Peptides and Polyclonal antibodies from SBL

- DNA Synthesis**
 - Oligo's 0.05 μ M–10 μ M
 - Labelled Fluorescein, Biotin 6-FAM, Hex, Tet et al
 - Purified HPLC/C18
- Peptide Synthesis**
 - Peptide Synthesis from 1–1000mgs
 - HPLC & Mass Spec Verification
 - HPLC Purification to 95% +
 - Conjugation to Carrier Proteins
 - Custom Polyclonal Antisera

For a Fast Efficient Service Reply Today

Severn Biotech Ltd, Unit 2, Park Lane, Stourport Road, Kidderminster, Worcs DY11 6TJ, UK
Tel: (01562) 825286 Fax: (01562) 825284
E-mail: 101626.2102@compuserve.com

READER ENQUIRY NO. 55